



**THIS PAGE BLANK (USPTO)**



636 CCGGAGAGCTGCTGAAGACTCTCAAGCTCAAGAGCCCTGCTACCCCAAGCTGGCGAC 695  
 QY 201 VALLIeGInGlyGlnAspPheValThraArgAgaGlieArgThrValValLeuAa 695  
 Db 696 CTGATCCAGGCTTGGCTGAGCGGAGGCTATTTCGAGAGCAGCTGGAGAGC 220  
 QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaIstYrile 755  
 Db 756 CTGCC-CCGCGTGGCCCTGCTCCAGGAGAGCCGACTGCTGCTCAAGTACATCAT 240  
 QY 240 uAsPLeuGlnArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 814  
 Db 815 GGAAGCTGGAGCGCTGGATCCAGCCGGCGCCCGGAGAGCCTTCAGCTGGGCTCCCTGG 260  
 QY 260 yAla-----GInGInGluProGlyLeuLeuArgValAlaGlyAspAsGlyIlePr 874  
 Db 875 GCGCCCTTGGCTGGCCAGAGAGCGCTGGGCTGCTCCGCTGGCTGGAGCGGCGATCGC 934  
 QY 277 GTrPserSerAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValas 934  
 Db 935 CTGGAGCCAGGAGAGAGAGAGGCTCCAGCCCTTGCAGACTTTCAGAAATCGTAGA 994  
 QY 296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGlnHisArgLeuValThra 316  
 Db 995 CATTAGCATTAAGCAGCGCCCGGCTGGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAG 1054  
 QY 316 ThrArgMetAspGlyHisIleLeuGlnAlaGluPheProGlyLeuProGluAlaLeu 1054  
 Db 1055 TACAG 336  
 QY 336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisIstYrPhe 1114  
 Db 1115 GTTCTGGCGCTGCTGGAG 356  
 QY 356 sLysGluValAlaProProArgLeuLeuGluGluGluAlaAspValCysHisArgProI 1174  
 Db 1175 CAAG 376  
 QY 376 ePheLeuAspPheAlaIleHisIstYrLeuLysAlaAlaGlySerLeuProGlyThrTyr 1234  
 Db 1235 CACTCTGGAGCTTGGCTATCAAGAGCTCAAGAGCTGGGCTGCTGCTGCTGCTGCTGCT 396  
 QY 396 eLeuArgAspSerProGlnAspPyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 1294  
 Db 1295 TCTCGGCGAG 416  
 QY 416 OLeuGlyProAspTyrIstYrGlyCysLeuIleArgGlnAspSerGlyAlaPheSerIe 1354  
 Db 1355 CTTTGGCTGCTGCTATTAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436  
 QY 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuAlaAlaGlyStrpAsnSe 1414  
 Db 1415 GGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 456  
 QY 456 rGlyLeuArgValAspGlyAlaAlaLeuThrIstYrSerCysAlaPheProGly 1474  
 Db 1475 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 476  
 QY 476 rGluYserSerAsnLeuIleValAlaArgArgIstYrCysAsnProAlaProAlaProGly 1534  
 Db 1535 AGAAAGCTCAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496  
 QY 496 sSerPro---SerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 1594  
 Db 1595 TCAGGCCCAATCCCAATACAGAGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 515  
 QY 515 rLeuGlnTyrPheIstYrAsnLeuGlyHisIstYrSerPheThrIstYrPheArgGlySer 1654  
 Db 1655 CTTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 535  
 QY 535 gArgGluValAlaAspGlyGluThrHisAspSerGlyValLeuLeuLysValMetAspSe 1714  
 Db 1715 CCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1774

QY 555 rArgHisArgAsnGlyMetGluSerPheLeuGlnAlaAlaSerLeuMetSerGlnValSe 575  
 Db 1775 CAAGCAG 575  
 QY 575 rTyrProHisIleValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValG 1834  
 Db 1835 GAG 595  
 QY 595 nGluPheValTyrLeuGlyAlaIleAspMetIstYrLeuArgGlyHisIstYrLeuValSe 1894  
 Db 1895 GGAATTTTACAGCTGGGAG 615  
 QY 615 rAlaSerTrpIstYrLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTrpIleuGln 1954  
 Db 1955 AGCCAGCTGGAG 635  
 QY 635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGly 2014  
 Db 2015 CAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655  
 QY 655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 2074  
 Db 2075 TGATGGAG 675  
 QY 675 rLeuGlnMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGlnAlaG 2134  
 Db 2135 CTTGAG 695  
 QY 695 nThrLeuCysLeuGlnAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheG 2194  
 Db 2195 GAGACTTAGCTTGAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 715  
 QY 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysIstYrLeuLysPheThrGlu 2254  
 Db 2255 TGCGCTGAG 735  
 QY 735 pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMe 2314  
 Db 2315 CCGGAG 755  
 QY 755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuAlaArgAspLeuAsnGly 2374  
 Db 2375 GGGCTTACAG 775  
 QY 775 uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 2434  
 Db 2435 CAGCTCTTACAG 795  
 QY 795 pGluLeuCysValAlaGlyAlaGlnLeuTyrAlaLysGlnAspProAlaIlePheGlnG 2494  
 Db 2495 TGGGCTGAG---AATGGTGGCCAGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815  
 QY 815 uArgHisIleLysIstYrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeu 835  
 Db 2552 GAG 855  
 QY 835 sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValGlyGlnLeuGln 2611  
 Db 2612 CCGCTTACAG 855  
 QY 855 sSerValProAspGlnGlnArgAspPheGlnArgGlyIleGlnIleLeuLysAlaLeu 2671  
 Db 2672 CAGCGGAG 875  
 QY 875 sSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeu 2731  
 Db 2732 CAGTGAATTTTATTTCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895  
 QY 895 gLeuValMetGlnTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgG 915  
 Db 2792 GCTGTCATGAG

**THIS PAGE BLANK (USPTO)**

[illegible]

COMMENT	27599-7295, USA
FEATURES	Witchuhn, B.A. Nature 370, 153-157, 1994.
source	Location/Qualifiers
gene	1. .3620 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="19" /map="19p12" /clone="SKBR-3" /cell_type="SKBR-3 breast cancer cells" 1. .3620 /gene="Jak3B" 96. .3380 /gene="jak3B" /function="tyrosine protein kinase" /codon_start="1" /product="JAK3B" /protein_id="AAC50226.1" /db_xref="GI:1039419" /translation="MAPSEETPLIPORCSLSLSTENGALHVLIPARGPQRLSPFS EGDILADLCVOAKKASGILPVYHSFLAFALTEDSCSPHIFSVEDASQVLLYRIR RSTYPMNGLEKRRFGIRLRDLAIIIDLPLYELHPONRSDLYSGRLPVGLSLKDG ESLAVYLDIRARMAEQOARRGELTKYSACLPSPSRDLOGLSTYTRIRITRRYR RALRYAACQADRSLMAKTYIMDLERLDPAEATETFGDGLAGHOGSLVAVAG GLIMATGDEVLDPFCDPEITVDISIKQARVPPEAGHRVLYTRIDNOITLEAFEGG LPEALSFVALYADGEVFLRTDSDQHFKEKQAPRLVEAEHCQPIITLDFAINLKLTGG GSRGYSYLRRSPDPDFSLTLCVQNPMLGADYIKGLRSPGTFLVIGSRSSLS RELATACWDGLAHVDAVLTSCCIPPKKSNLIVORSHSPSTSLVQPOQYOLO SOMFEAKIPMSDSVSYRHLVILGVCMDGSTMQOEFVILGADIDYLRKKDLPASKK SIPLEASIKMSQVSYRHLVILGVCMDGSTMQOEFVILGADIDYLRKKDLPASKK LQYKQALAYALNIYEDKGLPRGNVSARKVLLAREGADSPFFITLSDPVSRYVLSLE MLTRIPWAPLECELEAKQOTSLLEADKMGFGATVNEVSGVMPISALDPARKLOFYE RQQLPAPKWTLEALLIQOCMAVEVPQRFSAVLRDLNLSIDYELISDTPQALALAR LDHGMNGQALYACQDPTIFEEHRHLKYISQLKGNFSGVELCRYPDLGNTGALVAVKK LOHSGPDQDRDFOREIOTILKLSHDFIVKYGVSGVSGRQSLRLVMEITLPGCLARDF ORHARLADSLRLLYSOICKGMEYLRRCVHDLARNITLVESEAVKLTADGLA LLPLDKDYYVIREPGGSPITFWYAPESISDNIFFSROSDMVSQVLYELATFCMDSCSS SABRLRMGECERDVPALCRLLLELEBGRRLPAPACRAEVSAAGLASVSQSDMAGVS GRPGGA"
BASE COUNT	670 a 1152 c 1076 g 722 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3, 01e-273 Length: 3620
Score:	4597.00 Matches: 895
Percent Similarity:	86.80% Conservative: 65
Best Local Similarity:	80.92% Mismatches: 123
Query Match:	78.45% Indels: 24
DB:	9 Gaps: 7
US-09-397-967-16 (1-1099) x HSU31601 (1-3620)	
Qy	1 MetAlaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSer 20
Db	96 ATGGCACCTCCCAAGTAGAAGAAGACGCCCTCGATGCCCTCAACGACGCTTCCTG 155
Qy	21 SerGluAlaGlyAlaLeuLeuHisValLeuLeuProProArgArgGlyProGlyProGlnArg 40
Db	156 ACGGAGGCTGGAGCGCTCGATGCTGCTGCTCCGCCGCTCGGGGCCGCCGCCGCCGCC 215
Qy	41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db	216 CTATCTTCTCTCTTGGGAGCACCTGGCTGAGAGACCTGCTGCTGACGCTCCCAAGGCC 275
Qy	61 CysGlyIleLeuProValIleThrHisSerLeuPheAlaLeuAlaThrGlnAspPheSerCys 80
Db	276 AGCGGATCCTGCTGTGACCACTCCCTCTTGGCTCTGCGACACGGAGGACCTGCTCCG 335
Qy	81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db	336 TGGTTCCTCCCGGACCAACATCTTCGCGGAGGAGGACGACGACCAACATCTTCGCTGAC 395
Qy	101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120

THIS PAGE BLANK (USPTO)

五

Bosse, A., O'Shea, J.J. and Johnston, J.A.

THIS PAGE BLANK (USPTO)

E08798  
LOCUS E08798 165 bp RNA linear  
DEFINITION CDNA encoding novel human tyrosine kinase.  
ACCESSION E08798  
VERSION E08798.1 GI:2176910  
KEYWORDS JP 1995059569-A/1.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS Sakano, S.  
TITLE NOVEL TYROSINE KINASE PEPTIDE AND DNA CODING THE SAME  
JOURNAL Patent: JP 1995059569-A 1 07-MAR-1995.  
ASAHIT CHEM IND CO LTD  
OS Homo sapiens (human)  
PN JP 1995059569-A/1  
PE 07-MAR-1995  
25-AUG-1993 JP 1993210404  
SAKANO SEIJI

FEATURES  
source 1. 165  
FT /organism="Homo sapiens"  
FT /cell\_type="megakaryoblast"  
FT /cell\_line="U-7"  
FT /product="novel human tyrosine kinase"  
FT location/Qualifiers  
1. 165  
location/Qualifiers  
1. 165  
mat\_peptide  
/db\_xref="taxon:9606"  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 36 a 56 c 39 g 34 t  
ORIGIN

Alignment Scores:  
Pred. No.: 9 04e-09 Length: 165  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-397-967-15 (1-17) x E08798 (1-165)

OY 1 AlalysleuLeuProLeuAspTyrTyrValValArgGluProGly 17  
|||||  
Db 52 GCTAAGTCTGCTGCCGCTTGCACAAAGACTACTAGCTGTCGCCGACGACGAC 102

RESULT 2  
HSU08340 198 bp mRNA linear  
LOCUS Human clone NTK16 tyrosine kinase mRNA, partial cds  
DEFINITION U08340  
ACCESSION U08340.1 GI:473879  
VERSION 1  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS Fuortes, M.  
TITLE Tumor Necrosis Factor and Adhesion: Joint Control of Neutrophil  
JOURNAL Activation  
Thesis (1994) Cell Biology and Anatomy, Cornell University,  
Graduate School of Medical Sciences  
2 (bases 1 to 198)  
Fuortes, M.

TITLE Direct Submission  
JOURNAL Submitted (05-APR-1994) Michele Fuortes, Cell Biology and  
Anatomy/Medicine, Cornell University Medical College, 1300 York  
Avenue, New York, NY 10021, USA  
FEATURES  
source 1. 198  
location/Qualifiers  
1. 198  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="NTK16"  
/cell\_type="neutrophil"  
/tissue\_type="blood"  
CDS  
<1..>198  
/codon\_start=1  
/evidence=not\_experimental  
/product="tyrosine kinase"  
/protein\_id="AA17743.1"  
/db\_xref="GI:473880"  
/translation="DLARNLVSEAHVKAIDFGKLKLLPLDKYVYVREPGSPDF  
WYAPESLSDNFSRQSDVMSFG"

BASE COUNT 41 a 66 c 52 g 39 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.1e-08 Length: 198  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-397-967-15 (1-17) x HSU08340 (1-198)

OY 1 AlalysleuLeuProLeuAspTyrTyrValValArgGluProGly 17-  
|||||  
Db 67 GCTAAGTCTGCTGCCGCTTGCACAAAGACTACTAGCTGTCGCCGACGACGAC (117)

RESULT 3  
AX203080 448 bp DNA linear  
LOCUS Sequence 1 from Patent WO0152892.  
DEFINITION AX203080  
ACCESSION AX203080.1 GI:15392437  
VERSION 1  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Vasilos, G.  
TITLE Jak/stat pathway inhibitors and the uses thereof  
JOURNAL Patent: WO 0152892-A 1 26-JUL-2001;  
GENZYME CORPORATION (US)  
FEATURES  
source 1. 448  
location/Qualifiers  
1. 448  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 79 a 150 c 127 g 92 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.63e-08 Length: 448  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-397-967-15 (1-17) x AX203080 (1-448)

OY 1 AlalysleuLeuProLeuAspTyrTyrValValArgGluProGly 17  
|||||  
Db 306 GCTAAGTCTGCTGCCGCTTGCACAAAGACTACTAGCTGTCGCCGACGACGAC 356

RESULT 4  
LOCUS AX203080  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Thesis (1994) Cell Biology and Anatomy, Cornell University,  
Graduate School of Medical Sciences  
2 (bases 1 to 198)  
Fuortes, M.

THIS PAGE BLANK (USPTO)



GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 28, 2003, 17:00:25 ; Search time 6099 seconds  
(without alignments) 5244.136 Million cell updates/sec

Title: US-09-397-967-16  
Perfect score: 1099  
Sequence: 1 MAPSEETPLIPORSCSLSS.....RPAFATLSPQLDPLMGRPG 1099

Scoring table:  
OLIGO  
Xgapop 60.0, Ygapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 6.0, Deloxt 7.0  
Delop 6.0, Deloxt 7.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4106490

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Command line parameters:  
-MODEL=frame+pn.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool/US09397967/runat\_26042003.182926.11201/app.query.fasta.1.1287  
-DB=GenEmbl -QFMT=spool -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPT=0 -LOOPTXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human0.cdi -LIST=150  
-DOCALLIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=100 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HAPSPZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967@cgn2.1.14675@runat\_26042003.182926.11201 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEVTIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELOXT=7

Database:

GenEmbl: \*  
1: gb\_da: \*  
2: gb\_hg: \*  
3: gb\_in: \*  
4: gb\_ov: \*  
5: gb\_ov: \*  
6: gb\_ov: \*  
7: gb\_ov: \*  
8: gb\_ov: \*  
9: gb\_ov: \*  
10: gb\_ov: \*  
11: gb\_ov: \*  
12: gb\_ov: \*  
13: gb\_ov: \*  
14: gb\_ov: \*  
15: gb\_ov: \*  
16: gb\_ov: \*  
17: gb\_ov: \*  
18: gb\_ov: \*  
19: gb\_ov: \*  
20: gb\_ov: \*  
21: gb\_ov: \*  
22: gb\_ov: \*  
23: gb\_ov: \*  
24: gb\_ov: \*  
25: gb\_ov: \*  
26: gb\_ov: \*  
27: gb\_ov: \*  
28: gb\_ov: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1037	94.4	3723	10	MCSPYKIN
2	221	20.1	3528	10	MUSJAK3A
3	180	16.4	4016	10	MUSJAK3H
4	65	5.9	3778	10	RATJAK3
5	62	5.6	778	6	AX203087
6	62	5.6	2914	6	AX203084
7	62	5.6	3620	6	AX203084
8	62	5.6	4064	6	AX203081
9	62	5.6	4064	9	HSU09607
10	62	5.6	10326	10	MMU071201
11	62	5.6	13862	9	HSU70065
12	62	5.6	218208	2	AC073700
13	62	5.6	223734	2	AC073750
14	51	4.6	795	6	AX203086
15	50	4.5	159623	2	AC103459
16	40	3.6	448	6	AX203080
17	39	3.5	21323	9	AF513860
18	39	3.5	41006	9	AC007201
19	37	3.4	198	9	HSU08440
20	31	2.8	165	6	EB08798
21	30	2.7	3644	5	AF034576
22	29	2.6	811	6	AX203088
23	29	2.6	159623	2	AC103459
24	28	2.5	2881	9	BC028068
25	18	1.6	109	6	AX203082
26	18	1.6	726	6	AX203085
27	17	1.5	495	9	AF227557
28	17	1.5	466	9	HSU57096
29	16	1.5	1967	5	DRE5691
30	16	1.5	3435	6	AR129823
31	16	1.5	3495	6	AR048200
32	16	1.5	3495	6	AR068116
33	16	1.5	3495	6	AR071176
34	16	1.5	3495	6	AR177639
35	16	1.5	3495	6	162349
36	16	1.5	3495	6	187981
37	16	1.5	3500	9	AF001362
38	16	1.5	3629	6	AR136148
39	16	1.5	3629	6	AR143945
40	16	1.5	3629	6	192575
41	16	1.5	3731	10	MUSPTK
42	16	1.5	3731	10	RNU13396
43	16	1.5	3739	4	AF148993
44	16	1.5	4042	4	AB036337
45	16	1.5	4161	9	AF005216
46	16	1.5	4387	5	DRE5690
47	16	1.5	4599	4	AB036336
48	16	1.5	4812	4	AB006011
49	16	1.5	5117	6	AX354568
50	16	1.5	5117	9	AF058925

C	51	16	1.5	66974	2	AC119228	AC119228 Mus muscu
C	52	16	1.5	85671	2	AC096324_3	Continuation (4 of
C	53	16	1.5	17114	9	AL161450	AL161450 Human DNA
C	54	15	1.4	4744	10	AF173032	AF173032 Mus muscu
C	55	15	1.4	7134	10	AF052607	AF052607 Mus muscu
C	56	15	1.4	14907	5	AF090382	AF090382 Tetradon
C	57	15	1.4	172936	2	AC027154	AC027154 Mus muscu
C	58	14	1.3	8788	5	AF091238	AF091238 Tetradon
C	59	13	1.2	41	6	AR002620	AR002620 Sequence
C	60	13	1.2	45	6	AR009670	AR009670 Sequence
C	61	13	1.2	45	6	AR002619	AR002619 Sequence
C	62	13	1.2	45	6	AR099669	AR099669 Sequence
C	63	13	1.2	1592	8	AY078944	AY078944 Arabidops
C	64	13	1.2	16977	8	TF053213	TF053213 Tetradon f
C	65	13	1.2	107097	8	AC005727	AC005727 Arabidops
C	66	13	1.2	108393	8	AC005315	AC005315 Arabidops
C	67	12	1.1	151	6	144512	144512 Sequence 11
C	68	12	1.1	187	3	AB012713	AB012713 Anthocida
C	69	12	1.1	201	3	HYDTRKIND	HYDTRKIND Homo sapien
C	70	12	1.1	206	6	H006874511	H006874511 Sequence
C	71	12	1.1	214	6	AX203083	AX203083 Sequence
C	72	12	1.1	302	5	AB038987	AB038987 Gallus ga
C	73	12	1.1	373	5	CHRTYKDA	CHRTYKDA Chicken DNA
C	74	12	1.1	738	5	AR053289	AR053289 Sequence
C	75	12	1.1	790	10	RNC029010	RNC029010 Rattus no
C	76	12	1.1	877	5	DR089380	DR089380 Danio rerio
C	77	12	1.1	1128	10	MMSEK4	MMSEK4 Mus musculu
C	78	12	1.1	1153	5	AB025542	AB025542 Lampetra
C	79	12	1.1	1186	3	HYDTRKINA	HYDTRKINA Hydra vulga
C	80	12	1.1	1225	3	HYDTRKINA	HYDTRKINA Hydra vulga
C	81	12	1.1	1294	10	MMSEK2	MMSEK2 M. musculus
C	82	12	1.1	1398	6	AR053290	AR053290 Sequence
C	83	12	1.1	1473	5	DR05028	DR05028 Danio rer
C	84	12	1.1	1509	10	RNEK	RNEK Rat mRNA fo
C	85	12	1.1	1521	6	AR053291	AR053291 Sequence
C	86	12	1.1	1531	5	XLEPHRTK	XLEPHRTK Xenopus lae
C	87	12	1.1	1563	5	AB025541	AB025541 Epitretu
C	88	12	1.1	1568	9	HSURFRET	HSURFRET H. sapiens u
C	89	12	1.1	1582	5	AB025538	AB025538 Epitretu
C	90	12	1.1	1585	3	AB049595	AB049595 Ephydatia
C	91	12	1.1	1602	10	MMU18084	MMU18084 Mus musculu
C	92	12	1.1	1618	9	HUMERKLP	HUMERKLP Homo sapien
C	93	12	1.1	1619	3	AB049594	AB049594 Ephydatia
C	94	12	1.1	1619	9	HSMB01500	HSMB01500 Homo sapi
C	95	12	1.1	1635	10	MMSEK3	MMSEK3 M. musculus
C	96	12	1.1	1651	10	MMU05210	MMU05210 Mus musculu
C	97	12	1.1	1656	10	S77473	S77473 VNR-nonreco
C	98	12	1.1	1658	9	BC008655	BC008655 Homo sapi
C	99	12	1.1	1734	10	MUSNRK	MUSNRK Mus musculu
C	100	12	1.1	1744	9	S71669	S71669 Lsk-leukocy
C	101	12	1.1	1804	9	HSMB02388	HSMB02388 Homo sapi
C	102	12	1.1	1829	3	SLSRK1	SLSRK1 X61601 S. lacustris
C	103	12	1.1	1838	10	RATBATK	RATBATK Rattus norv
C	104	12	1.1	1867	9	HUMREMRI	HUMREMRI Homo sapien
C	105	12	1.1	1911	10	MUSCTR	MUSCTR Mouse mRNA
C	106	12	1.1	1942	6	AR053292	AR053292 Sequence
C	107	12	1.1	1942	6	E10345	E10345 CDNA encodi
C	108	12	1.1	1968	6	HSHTYTK	HSHTYTK H. sapiens H
C	109	12	1.1	1987	6	AR084686	AR084686 Sequence
C	110	12	1.1	1987	6	AR204704	AR204704 Sequence
C	111	12	1.1	1987	9	HUMMATK	HUMMATK Human tyros
C	112	12	1.1	2000	6	ARI79662	ARI79662 Sequence
C	113	12	1.1	2073	9	AK055395	AK055395 Homo sapi
C	114	12	1.1	2112	10	MMEBK	MMEBK M. musculus
C	115	12	1.1	2170	6	AX015415	AX015415 Sequence
C	116	12	1.1	2180	9	BC000114	BC000114 Homo sapi
C	117	12	1.1	2180	9	BC003109	BC003109 Homo sapi
C	118	12	1.1	2219	9	BC003072	BC003072 Homo sapi
C	119	12	1.1	2318	3	AB006566	AB006566 Ephydatia
C	120	12	1.1	2398	9	HSRTKEPH	HSRTKEPH H. sapiens m
C	121	12	1.1	2485	9	HUMTYKRET	HUMTYKRET Human ret m
C	122	12	1.1	2578	5	XLA236868	XLA236868 Xenopus 1
C	123	12	1.1	2592	5	XLA236867	XLA236867 Xenopus 1

124	12	1.1	2640	5	DR05030	DR05030 Danio rer
125	12	1.1	2720	9	HSEPHAT5	HSEPHAT5 Homo sapi
126	12	1.1	2735	3	AB006565	AB006565 Ephydatia
127	12	1.1	2737	3	AB005305	AB005305 Homo sapi
128	12	1.1	2784	6	AX481390	AX481390 Sequence
129	12	1.1	2820	5	GGECK8A	GGECK8A Sequence
130	12	1.1	2820	6	115002	115002 Sequence 5
131	12	1.1	2839	6	AF037333	AF037333 Homo sapi
132	12	1.1	2872	9	HSRETTT	HSRETTT Homo sapi
133	12	1.1	2931	5	DR05029	DR05029 Danio rer
134	12	1.1	2931	10	AF131197	AF131197 Mus muscu
135	12	1.1	2955	6	AX481387	AX481387 Sequence
136	12	1.1	2955	6	AX481407	AX481407 Sequence
137	12	1.1	2962	6	AR084933	AR084933 Sequence
138	12	1.1	2962	6	AR084949	AR084949 Sequence
139	12	1.1	2962	6	HUMPRKA	HUMPRKA Homo sapien
140	12	1.1	2964	6	AX481410	AX481410 Sequence
141	12	1.1	2982	6	A44405	A44405 Sequence 2
142	12	1.1	2982	6	AR038071	AR038071 Sequence
143	12	1.1	2982	6	AR172554	AR172554 Sequence
144	12	1.1	2982	6	119546	119546 Sequence 2
145	12	1.1	2997	6	AX481391	AX481391 Sequence
146	12	1.1	2997	6	AX481409	AX481409 Sequence
147	12	1.1	3042	5	XLSK1	XLSK1 X. laevis mr
148	12	1.1	3042	6	AX481388	AX481388 Sequence
149	12	1.1	3055	5	DR05026	DR05026 Danio rer
150	12	1.1	3104	10	MUSNRK	MUSNRK Mus musculu

## ALIGNMENTS

RESULT 1  
MUSPYTKIN 3723 bp mRNA linear ROD 21-SEP-1994  
LOCUS  
Mouse protein tyrosine kinase (jak3) mRNA, complete cds.

DEFINITION  
L32955.1 GI:529238  
Interleukin 4; Janus kinase; protein tyrosine kinase.

VERSION  
L32955.1 GI:529238  
Interleukin 4; Janus kinase; protein tyrosine kinase.

KEYWORDS  
Mus musculus (strain BALB/c, sub-species domesticus) cDNA to mRNA.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
1 (bases 1 to 3723)  
Liu, E.T., and Lin, N.  
Involvement of the Jak-3 Janus kinase in signalling by interleukins 2 and 4 in lymphoid and myeloid cells

AUTHORS  
Witthuhn, B.A., Silvenoinen, O., Miura, O., Lai, K.S., Cwik, C.,  
Nature 370 (6485), 153-157 (1994)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
8022486  
This entry has been reported under the accession number L329555 in the Nature article.

FEATURES  
source  
location/Qualifiers  
1..3723  
/organism="Mus musculus"  
/strain="BALB/c"  
/sub-species="domesticus"  
/db\_xref="taxon:10090"  
/cell\_line="Abelson transformed pre-B-cell line 22D6"  
1..3723  
/cell\_type="lymphoid and myeloid"  
1..3723  
/gene="jak3"  
1..12  
/gene="jak3"  
13..3312  
/gene="jak3"  
/standard\_name="Janus kinase, mujak3"  
/codon\_start=1  
/evidence="experimental"  
/product="protein tyrosine kinase"  
/protein\_id="AA021565.1"  
/db\_xref="GI:529239"

gene  
5'UTR  
CDS

/translation="MAPSEETPLIPORSCSLSSSEAGALLVLPKPGPPORLSF  
EGDVLADICVRAKACALIPVHSILFLEDEDSCEPSPSHICIDVPOVIVYL  
RFPYPMFGLFETCHREGRRLKRDLSAIIIDLIVLELPHORHSDLVSGRLPYGLSKMEG  
EFLSLAVDLAOMARDOARORBELLTAKIPSLRDVIOGONVYTRIRTRITY  
LALLPCRLPGRPALMAKYILDERLHPATETTFVGLCAOEFGLLRVADNBI  
PMSNDELQOTFCDFPEIADVSIINQARVPAGHRVITVIMGHILAEFLP  
LSFVALVDGYFRLICDSRHYECKEVPARLLEEADVCHPITLDEAIIHKLKAAGLIP  
GTYLARSPODOSFLTACVOPFLGQCLIROPSGSAFVGLGSDPSHRL  
AACNSCLRDGALVYITSCAPPRKSNLIYVRRCALNAPAGCSPPSCALITLSF  
HTITPDSLEHENTLHGSEFTKIFRGSREVYDGTTHSEVLKYMDSRHRNCHSEFL  
AASMSQVSTPHVLHLHGVCWAGDSINQEVVIGALIDMTLRKRGHLVLSWKLQVTK  
QLAVALNYLEDKGLPHGNVSAKRYLARREGDGNPPIKILSDPVSPTVLSLEMLTOR  
IPVAYPELOEAQTLCEADKMGFATTEVEFORGPAHITSLEPAKRLKEDEMGOLP  
ALKWTELGLITOCMAVDPGRRSPFRAILRDNLNITSDYELSDPTPGIIPSPDELC  
VAGNOLACODDPAIFEEERHLKYISLKGNGSVELCRVDLGNQTPGLAVAKOLOHS  
VPDQORPORIOTIKALHSDFIYKRGVSGPRLGROSLRIMEVLPESGCRDLLOHR  
GLHTRDLRFRAWQICKMEYIGARCRHDAANILYVESAHYKIKADFGAKILPLIG  
KDYVVEREGOSPIFWYAPESLSDNITSKSDVSEGVLYELFTYCDKSCSPAEEL  
RMGPEREGPPLRLLELAEGRRLPPPTCPTEVQELMDLQMAPEBHDPAFATLSF  
QLDPLMGRC"  
13..3309  
/gene="jak3"  
/product="protein tyrosine kinase"  
/standard\_name="Janus kinase, mjak3"  
/evidence=experimental  
3313..3723  
/gene="jak3"  
3723  
/gene="jak3"  
polyA\_site  
BASE COUNT 711 a 1162 c 1079 g 771 t  
ORIGIN

## Alignment Scores:

Pred. No.:	0	Length:	3723
Score:	1037.00	Matches:	1098
Percent Similarity:	99.82%	Conservative:	0
Best Local Similarity:	99.82%	Mismatches:	1
Query Match:	94.36%	Indels:	2
DB:	10	Gaps:	0

US-09-397-967-16 (1-1099) x MUSPTYKIN (1-3723)

QY 1 MetAlaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
DB 13 ATGGACCTCCCAAGTGGAGAGACACTCGATCCCTCACGCTCTGCGACCTTCATCC 72  
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40  
DB 73 TCAGAGGAGGAGACCTTGATGCTCTCTCCCGGGGAGACTGGGCTCCCAAGCA 132  
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaIleVala 60  
DB 133 TTGTCAATCTCTTTGGGAGACTACTTGCGTGAAGATTATGTGTGCGACCTGCCAAGGCC 192  
QY 61 CysGly-IleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerC 80  
DB 193 TGTGCG-TATTCGCGCTGTTATCATTCGCTTTGCGCTGCGCACAGGACTTCTCTTG 251  
QY 80 STrpPheProProSerHisIlePheCysTleGluAspValAspThrGlnValLeuValT 100  
DB 252 CTGGTTTCCCAAGCCACATCTTCTGCATAGAGAGCGTGGACACTCAAGTCTTGCTTA 311  
QY 100 rArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLe 120  
DB 312 CAGGCTACGCTTTTATTTCCCTGACGTGTTGGCGTGGAGACATGTCACGCGTTGGCGT 371  
QY 120 uArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaG 140  
DB 372 CGCAAGAGATTGGACAGGACATCTTGACTTACATGTTTAAAGACATCTCTTGCTCA 431  
QY 140 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGlnGlnG 160  
DB 432 GCACCGCACTGCTGCTGAGCGCTCCCGGTGGGCTTAAAGCATGAGAGAGGAG 491

QY 160 YGIuPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
DB 492 AGAGTCTCTGAGCTCGCGCTGCTGAGACTTGGCCAGATGGCTGTGAGAGCGGCCAGCG 551  
QY 180 gProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAs 200  
DB 552 CCCAGAGAGCTGCTGAAGAGCGTCAGTTACAAAGCTGTCTCGCGCCACCGCTCGCGCA 611  
QY 200 pValIleGlnIleGlnAsnPheValThrArgArgArgIleArgArgTrpThrValValLeuAl 220  
DB 612 TGTGATCCAGGAGCCAGAACTTGTCAGCAGAGCGGCATCCGACAGAGCGTGGTCTGGC 671  
QY 220 aLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLe 240  
DB 672 GCTGTGCGCGTGTGCTGCTGCGCAGCGCGAACCGTACAGCGCTGCAAGCATATATCT 731  
QY 240 uAspLeuGluArgLeuHisIleProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260  
DB 732 GGACCTTGAGGCGGCTACATCCAGCGCGCCACACCGACCTTCCGTGGGGCTCCCGGG 791  
QY 260 YAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSe 280  
DB 792 CCGCCAGAGAGAGCGCGGGCTTCTGCTGTCGCGAGGGAGCAACGCAATCCCTTGAGCTC 851  
QY 280 rAsnAspLeuLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAs 300  
DB 852 GAACGACGAGCTTTCAGACCTTCTGTGACTTCCGGAATCCGATGTCAGCATCA 911  
QY 300 nGlnAlaProArgValGlyProAlaGlyGlnHisArgLeuValThrValThrArgMetAs 320  
DB 912 TCAGGCGCCACGTGTGGTCCGCGAGGAGACCGCGCTGGTGCATGTCACAGGATGGA 971  
QY 320 pGlyHisIleLeuGluGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLe 340  
DB 972 CGGCCACATCTCTGGAAGCGGATTTCCGGGGCTGCTAGGGCGCTCTTGTGGCCCT 1031  
QY 340 uValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAl 360  
DB 1032 CGTGATGGGTACTTCCCGCTGATCTGCGACCTCCAGGCACTTATTTCTCCAAAGAGCTGGC 1091  
QY 360 aProProArgLeuLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPh 380  
DB 1092 GCCGCCAAGGCTGCTGGAGAGAGAGCGGACTGTGGCTGACCATGACCTTATGACTT 1151  
QY 380 eAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSe 400  
DB 1152 TGCATCCACAAAGCTGAAGGCGGTGGCTCCCTCCAGGACCTATATTTCTCGCGCGAG 1211  
QY 400 rProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAs 420  
DB 1212 CCGGAGAGACTATGACAGCTTCTTCTTAACCCCTGCGCCAGACTCCCTTGGGCCCA 1271  
QY 420 pTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSe 440  
DB 1272 CTACAAAGGCTCTCATCGCAGAGACCCACAGGGGCTTCTCTCCCTGGTGGCTCAG 1331  
QY 440 rGlnProHisArgSerLeuArgGlyIleuLeuAlaAlaCysTrpAsnSerGlyLeuArgVa 460  
DB 1332 CCAGGCCCAAGAGCTCGGGAGCTGTTGCAGGCTCTGGATTTGGGCTGGAGAT 1391  
QY 460 lAspGlyAlaAlaLeuTyrLeuThrSerCysAlaProArgProLysGlnLysSerAs 480  
DB 1392 AGACGCTGCTGCCCTGTACTTAACATCCGTGCGCTCCACAGCCCAAGAAAGTCCA 1451  
QY 480 nLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCy 500  
DB 1452 TTTGATCTGTGGTCCAGAGGGGCTGCAACCCCGCGCTGCGCTGGTCTCCCGTCCG 1511  
QY 500 sCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisG 520  
DB 1512 CTGTGCGCTGACAGAGCTAGCTTCCACACAAATTCACAGAGAGCTGAGTGGAGAGA 1571

QY 520 uasnleuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValAs 540  
 |||||  
 Db 1572 GAACCTGGGTACAGCTCTTTTACCAAGATCTCCGTGGCAGCAGCGGAGGCTGTGA 1631  
 QY 540 pGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCy 560  
 |||||  
 Db 1632 TGGGAGACATGACTCTGGAGAGTCTCTCAAGTCTATGACTCCAGACATCCGAACTG 1691  
 QY 560 smetGluSerPheLeuGluValAlaSerLeuMetSerGluValSerTyrProHisLeuVal 580  
 |||||  
 Db 1692 CATGAGCTCTTTCTGGAGAGCCGCAAGCTTATGACCAAGATCTTACCCGACCTGCT 1751  
 QY 580 lleuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyrIle 600  
 |||||  
 Db 1752 GTTACTGCACGCGCTGCTCATGTGGTGAGACAGATCATGTGGCAGAAATTGGTATCT 1811  
 QY 600 uGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrPylsLe 620  
 |||||  
 Db 1812 AGGAGCAATTGACATGTACCTGGCCCAAGCCGTGGCCACCTGCTGTACGCCACTGGAACT 1871  
 QY 620 uGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrIleGlnLysAspLysGlyLeuProH 640  
 |||||  
 Db 1872 GCAGGTGACCAAGCAGCAGCTGGCATATGCCCTTAATCTTGGAGAGCAAAAGGCTTCTCA 1931  
 QY 640 sGlyAsnValSerAlaArgLysValLeuLeuAlaArgGlnGlyLysAspGlyAsnProPr 660  
 |||||  
 Db 1932 CGGCAAGCTCTCACAGCAGGAGGAGTGTCTGCTGCTGAGGGGGTGTATGGAAATCCACC 1991  
 QY 660 oPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuTh 680  
 |||||  
 Db 1992 TTTATTAAGCTGAGTGTCTGTGGTGTGATCCACAGTGTGTGAGCCCTGGAAATGCTCAC 2051  
 QY 680 rAspArgIleProThrPylAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuG 700  
 |||||  
 Db 2052 CGACAGAAATACCTGGGGGGCCCGCAATGTCCTCAGAGAGCTCAGACACTCTCTTGA 2111  
 QY 700 uAlaAspLysTyrPylGlyPheGlyAlaThrThrTyrGluValPheGlnArgGlyProAlaH 720  
 |||||  
 Db 2112 GGGGCAAGAGGGGGCTTTGGAGCCACCACCTGGAGAGTGTCCCAAGGGGAGCCGCCCA 2171  
 QY 720 sIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGlnAspGlnGlyGlnLeuPr 740  
 |||||  
 Db 2172 CATACCTCGCTGGAGGCCCGCAAAAAGCTTGAAGTTCTATGAGAGCCAGGAGACAGCTGCC 2231  
 QY 740 oAlaLeuLysTyrThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProG 760  
 |||||  
 Db 2232 CGCTCTCAAAATGACAGAACTGGGGAGCTTATCACACAGATGATGGATGATCTCTGG 2291  
 QY 760 yArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTy 780  
 |||||  
 Db 2292 CCGGCGCCCTCTTCGAGCTATCTCAGAGACTCAACGCGCTCATTTACATCAGATTGA 2351  
 QY 780 rGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAl 800  
 |||||  
 Db 2352 CGAGCTCTCTCAGAGCCCAACACTGGCATCCGAGTCTTCGAGATGAGCTGTCTGAGC 2411  
 QY 800 aGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLysTy 820  
 |||||  
 Db 2412 TGGGGCCCACTATCTCCCTGCCAGGAGCCCGCCATTTGAGAGAGAGACACTTAAGTA 2471  
 QY 820 rIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLe 840  
 |||||  
 Db 2472 CATCTCTTCTGGGCAAGGAGCACTTGGAGGCTGGAGCTGGCGCTATGACCCCT 2531  
 QY 840 uGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspG 860  
 |||||  
 Db 2532 GGGGAGACAAATACGGGAGCCCTGGTGGCAGTGAAMAACACTACAGACAGCGTCCAGACA 2591  
 QY 860 nGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880  
 |||||  
 Db 2592 GCAGAGGAGACTTCCAGCGGAGATTCAGATCTTAAGGCTCTGACACAGCACTTCATCT 2651  
 QY 880 lLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTy 900

Db 2652 CAAGTACCGGGGAGTACACTATGGCCAGAGTCCCAAGACCTGCTTGTGGATGAGATA 2711  
 QY 900 rLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGlyLeuHisThrAspAr 920  
 |||||  
 Db 2712 CTGGCCCAAGCGGCTGCCGAGAGACTTGTCTCAGCGCATGCGGCGTCCAGACCGACCG 2771  
 QY 920 gLeuLeuLeuPheAlaTyrPylIleCysLysGlyMetGluTyrLeuGlnGlyAlaArgArgCy 940  
 |||||  
 Db 2772 CTTACTGCTGTGCTGGCAGATCTGCAGAGGATGGATGAGTACCTGGGGTGGCCCGCTG 2831  
 QY 940 sValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValLysI 960  
 |||||  
 Db 2832 CGTACACCGCTACCTGGGTGGCCCAACATTTGTGTGAGAGCGAGGCTCATGTGAGAT 2891  
 QY 960 eAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgG 980  
 |||||  
 Db 2892 CGCGGACTTCGCGCTCGCTAGCTGTGCTGCCCTGGAAAGACTACTACGTGTGCCGGA 2951  
 QY 980 uProGlyGlnSerProIlePheTyrPylAlaProGluSerLeuSerAspAsnIlePhe 1000  
 |||||  
 Db 2952 GCTTGGCCCAAGCCCATCTTTTGTGTATGCCGGAGTCCCTATCTGACACATCTTCTC 3011  
 QY 1000 rArgGlnSerAspValTyrSerPheGlyValValLeuTyrGlnLeuPheThrTyrCysAs 1020  
 |||||  
 Db 3012 CCGCCAACTGACGTGTGAGATTTCCGAGTGTGTGACAGACTTTCACCTACTGCGA 3071  
 QY 1020 pLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyPr 1040  
 |||||  
 Db 3072 CAAAGCTGAGCCCATCTCCGTGAGTCTCCGCGATATGGGGCTTGAAGAGACC 3131  
 QY 1040 oProLeuCysArgLeuGlnLeuLeuAlaGlnGlyArgArgLeuProPropProTh 1060  
 |||||  
 Db 3132 CCGCTCTGCGCGCTCTCGAGAGCTGTGAGAGAGGCGGAGGCGGAGCTCCACACTCCAC 3191  
 QY 1060 rCysProThrGlnValGlnGlnLeuMetGlnLeuCysTyrPylAlaProGluProHisAspAr 1080  
 |||||  
 Db 3192 CTGGCCCAAGAGGTTGAGGAGCTCATGACAGCTGTGTGGCGGCCCGCAGCGCAGCAGC 3251  
 QY 1080 gProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTyrPylArgGlyArgProGly 1099  
 |||||  
 Db 3252 GCCAGCTTGGCGACCTGTGAGCCGCCAGCTGAGACCCGCTGTGGCTGGAAGACCCGGA 3309  
 RESULT 2  
 MUSJAK3A 3528 bp mRNA Linear ROD 24-Apr-1996  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus  
 thymus cDNA to mRNA.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 3528)  
 AUTHORS  
 Gurniak,C.B. and Berg,L.J.  
 TITLE  
 Murine JAK3 is preferentially expressed in hematopoietic tissues  
 and lymphocyte precursor cells  
 JOURNAL  
 BLOOD 87 (8), 3151-3160 (1996)  
 MEDLINE  
 96184772  
 PUBMED  
 8603529  
 FEATURES  
 source  
 location/Qualifiers  
 1..3528  
 /organism="Mus musculus"  
 /strain="BALB/c x 129 F2"  
 /db\_xref="taxon:10090"  
 /clone="clone-s12-18"  
 /cell\_type="T cell precursor"  
 /tissue\_type="thymus"  
 /dev\_stage="neonate"  
 1..3528  
 /gene="JAK3"  
 gene

```

5'UTR
1. 71
/gene="JAK3"
72. .3374
/gene="JAK3"
/codon_start=1
/protein_id="AAC42085.1"
/db_xref="GI:1019910"
/translation="MAPSEETPLIPORCSLSSEAGALHVLPPRGPPORRISFSE
EGDYLAEDLCVRKAKAGCILPVHSLEFLATEDESCMPSPHIFCIEEDVQVLYRL
REVPEDMELETCRHRGFLKDLTSAIDLHLEHLEFADHSDLVSGRLPVGLSKMEOG
EFLSLAVIDLADMADEORARPEGLIKTVSYACILPSLRDIOGONFPTRRIRRTVY
LALRRVACQADRALAMAKYILDELRHPATTEFRFRGLGAEPEERLLEFACDNGI
SWSSGDELFOFDEPVIDVSTKQARVPAGHRIPTITMDGHLELEFELPE
ALSFVALVDGYFRLICDSRHAYCEKAVAPRLLAEELCHGPIILDEFAIHKLAGSL
PGTYILRRSPDYDSFLTLACVQTPLEGPDYLGCLLRODPSGAFSLVGLDSQPHRSIREL
LAACMSRLRVDAALNLTSCAPRPKSNLIYVRRCCTPAPAPGSCPSGCALTOIS
FHTIPDSELEHENGHSFTKIFGRREYVDDGETHDESVLLKVMDSRHRNCKMSEI
EASLMSQVSPHLYLLHGNMAGDSIMVQFVYLGAIDMYLKRKRGHVSASMKLOVT
KQLYALNTLEDKCLPHGNVSARKVLLAREGGDGNPITKLSDRGVSTVLSLEMLTD
RIPWAPRCLQEAQTLCELDKMGFATTWVEFSQGNHITISLEPAKLLKYEEOGL
PALMWTLAGLITQCMAYDPGRSPFRALIRDLNGLITSDVELSDPTGPIPSDEL
CGAQLYACODPAIFEERHLKYISLKGNGFSVELCYDPLGDMTGPVLAVKLOHS
GPDQDRFOREIOLIKALHSDFIYKRGVSYGPGROSRLVMEYLPSCGLDPLORHR
ARLTDRLLEFAMQICKMEYLGAARCYRDLARNLITSEAHYKIKDDECLATLPL
GKDYVVAPEQCSPIFWTAPESLSDNITSROSDVASEVVIYELPTTDKSKCSBAER
LRMGPEREGRPLCRLELLAEGRRLPPTCPTEVQELMOLCVAPSPDHPAREGTLIS
POLDALMRGRG"
3'UTR
3375. .3528
/gene="JAK3"
3517. .3522
polyA_signal
/gene="JAK3"
BASE COUNT 648 a 1104 c 1036 g 740 t
ORIGIN

Alignment Scores:
Pred. No.: 7 67e-205 Length: 3528
Score: 221.00 Matches: 1060
Percent Similarity: 96.89% Conservative: 0
Best Local Similarity: 96.89% Mismatches: 15
Query Match: 20.11% Indels: 34
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x MUSTAK3A (1-3528)
QY 1 MetAlaProPheSerGluGluThrProLeuIleProGlnIleArgSerCysSerLeuSerSer 20
DB 72 ATGGGACCTCCCAAGTAGAGAGACACTGTGATCCCTGACGCTCTTGACGCTCTCATCC 131
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProAlaGlyProGlyProProGlnArg 40
DB 132 TCAGAGGGGAGGAGCCCTGTGATGTCTCTCCCGGGGACCTGGGCTCCGCCAGCGA 191
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaVala 60
DB 192 TTGTCTATTTCTTTTGGGAGCTACTTGGCTGAGGATTTATGTGTGGAGACTGCCAAGGCC 251
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGlnAspPheSerCys 80
DB 252 TGTGGGATCTGCTGCTTATCATTCGCTTTCGCTCTGGCCACGTGAGGACTTCTCTTGC 311
QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
DB 312 TGGTTTCCCCCAAGCACATCTTTCGATVAGAGAGCTGGACACTCACTCACTTGGCTCTAC 371
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
DB 372 AGGCTACGCTTTTATTTTCCCTGACTGGTTTGGGCTGGAGACATGTACCGCTTTGGGCTG 431
QY 121 ArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
DB 432 CCCAAGAATTTGACGACGAGTCCATCTTGTACATGATTTTGAACATCTCTTTGGCTAG 491
QY 141 HisArgSerAspLeuValSerCylArgLeuProValGlyLeuSerMetGlyGluGlnGly 160

```

```

DB 492 CACCCAGTACAGCTGTGAGTGGGCGCTCCGCGTGGGCTTACGATGAGAGAGAGCGA 551
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnIleArg 180
DB 552 GAGTTCTCGAGCTGTGCGCTGCTGAGCTTGGCCAGATGGCTGTGAGAGAGCGCCAGCG 611
QY 181 ProGlyGluLeuLeuHisThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
DB 612 CCAGGAGAGCTGCTGAAGAGCGGTCAAGTTCAAAGCGTGTCCGCGCCAGCGCTGGCGCAT 671
QY 201 ValIleGlnGlyGlnAspPheValThrArgArgGlyIleArgArgThrValValLeuAla 220
DB 672 GTGATCCAGGGCCAAACTTCGTGACACCCAGCGCGATCCGAGAGACCGTGGCTTGGGG 731
QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240
DB 732 CTGGC-CCGCTGTGTGTCGCTGCGCAGGCGCAGCGCTCAAGCGCTATATTTCT 790
QY 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
DB 791 GGACCTGGAGCGGATACATCCAGCGGCGCACCCGAGACCTTCGCTGGGCTCCGCGG 850
QY 260 ValArgGlnGluArgProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePro-TrpSerS 280
DB 851 CCGCCAGGAGGAGCGGGGCTTCTGCTGTGGCGGGGCAACGGCATCTC-CTGGAGCT 909
QY 280 eFAsn-Asp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSer 298
DB 910 CCGG-GGACCGAGAGCTTTTCCAGACCTTCTGTGACTTTCCGGMAATCTGCATGTACAC 968
QY 299 IleAsn-GlnAlaProArgValAlaGlyProAlaGlyGlnHisArgLeuValThrValThrAr 318
DB 969 ATCAA-GCAGGCCCCACGTTGGTGGTCCGCGAGGAGGACCGCGCTGCTCATCTGCACAG 1027
QY 318 GMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheVal 338
DB 1028 GATGAGCGGCGCACATCTGGAAGCGAGTTCCGGGGCGCTGTGAGCGCTGCTTTCT 1087
QY 338 lAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysG 358
DB 1088 GCGCCCTGCGATGGGTACTTCCGCTGATCTGCACACTCAGAGCATATTCTTGCAGAGA 1147
QY 358 uValAlaProProArgLeuGluGluGluAlaAspVal--CysHisGlyProIleThr 377
DB 1148 GGTGGCGCGCCACGCGCTCTGAGAGGAGAGCGGAG-GC-TGTGCATGTGACCAATCAGC 1205
QY 378 LeuAspPheAlaIleHisTyrLeuHisValAlaGlySerLeuProGlyThrTyrIleLeu 397
DB 1206 TTAGACTTTGCCATCCACAAGCTGAAGGCGGCTGCTCCCTCCAGGACCATATATCTTC 1265
QY 398 ArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeu 417
DB 1266 CGCGGAGCGCGCAGAGACATATGACAGCTTCTTCTTACGCGCTGCTCCAGACTCTCTT 1325
QY 418 GlyProAspTyrGlyGlyCysLeuIleArgGlnAspProSerGlyValaPheSerLeuVal 437
DB 1326 GCGCCCGAGTACAGAGGCGCTCATCCGCGAGAGACCCAGGCGGCTTCTTCCCGCTG 1385
QY 438 GlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGly 457
DB 1386 GGCCTTCAGCCACCCACAGGAAGCTCGCGGAGCTCTTGGACGCTGTGGAATCTTGGG 1445
QY 458 LeuArgValAspGlyAlaAlaLeuTyr-LeuThrSerCysCysAlaProArgProLysG 477
DB 1446 CTGCGAGTGAAGGCTGCTCCCTGAA-CTTAACATCTCTCTGCGTCCCAAGACCAAGA 1504
QY 477 uLysSerAsnLeuIleValValArgArgGlyCysAsn-ProAlaProAlaProGlyCysS 497
DB 1505 AAAGTTCATTTGATCGTGTGCGAAGGGGCTGAC-CCCGCGGCTGCGCCCTGCGTCT 1563
QY 497 eTrProSerCysAsnAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuG 517

```

Db 1564 CCCCCTCCTGCTGTCGCTGACACAGCTGAGCTTCACACAAATTCACAGGACAGCTGG 1623  
 QY 517 lUTPRHISGLIASLLEUGLYHISGLYSERPHETHLYSILEPHEARGGLYSER-ArgArg 536  
 Db 1624 AGTGACAGAGAACCTGGGTGCTGACGGTCTTTTACCAAGATCTTCCTGGCCG-CAGGGGG 1682  
 QY 537 GLUVALVALASPGLYGLUTHRHISASPSERGLUVALLEULEULYSVALMELASPSERArg 556  
 Db 1683 GAGGTCCGTGATGGTGGAGACACATGACTCGGAAGTCCCTCGAAGGTGATGACTCCAGA 1742  
 QY 557 HISARGANCSYMETGLUSER-PHELEUGLUALAASERLEUMETSERGLINVALSERTYR 576  
 Db 1743 CATGGAACCTCATGAGATCTTTCTGGAACCCCAAGCTTGATGAGCCAGATATCTTAC 1802  
 QY 577 PROHISLEUVALLEULEUHHISGLYVALCYSMETLALAGLYASPSERILEMETVALGLINGLU 596  
 Db 1803 CCGAGACCTGGGTGTTACTGTCACAGGCGTGTGCAATGCTGGAGACACATCATGATGGAGAA 1862  
 QY 597 PHEVALTYRLEUGLYALALIASPMETTYRLEUARGLYSARGLYHISLEUVALSERALA 616  
 Db 1863 TTTGTGATCTAGAGAGCAATTGACATGACTGCGCAAGCTGGCCACCTGCTGCACACC 1922  
 QY 617 SERTPRYSLEUGLINVALTHLYSGLEUVALATYRVALALEUASNTYRLEUGLIASPLYS 636  
 Db 1923 AGCTGGAACACTGAGGTGACCAACAGCTGCAATGCTTAACCTAGTGGAGACAAA 1982  
 QY 637 GLYLEUPROHISGLYASVALSERALARGLYSVALLEULEUVALARGGLUGLYASP 656  
 Db 1983 GGCCTTCTCAGCGCACAGCTTCAGCACAGGAGAGTGCCTCGCTCGTGGAGGGGGTAT 2042  
 QY 657 GLYASNPPOPHELLEULYSLEUSERASPPROGLIYVALSERPROTHVALLEUSERLEU 676  
 Db 2043 GGGAACTCACCTTCATTAACCTAGTATGATCTGATGCTGATCCACAGTCTGAGCCTG 2102  
 QY 677 GLUWETLEUTHRRASPARGLILEPOTRYVALALAPROGLUCYSLEUNGILUALAGLINTHR 696  
 Db 2103 GAAATGCTCACCGACAGAAATACCTGGGTGGCCCCGAAATGCTTCACAGAGGCTCAGACA 2162  
 QY 697 LEUCYSLEUGLUALIASPLYSSTRPGLYPHEGLYVALATHRRTPRPGLYVALPHEGLINARG 716  
 Db 2163 CTCGCTTGAGAGGCTGACAAATGGGGCTTTGGACCCACACAGTGGAGAGTGT-CAGGG 2221  
 QY 717 -GLYPROALAHISILETHRSERLEUGLUPROALALYSLEULYSPHYTYRGLIUSPGL 736  
 Db 2222 GGGACCCGCCACATCACTGCTGAGACCCGCCAAAAGCTGAAGTCTATGAGAGACA 2281  
 QY 736 NGLYGLEUPROALALEUSTRPPTHGLULEUALAGLYLEUILLIETHRGINCYSMETAL 756  
 Db 2282 GGGACACTGCGCCCTCTCAATGAGACAGACTGGCGGACTTATCACACAGTGCATGCG 2341  
 QY 756 ATYASPPROGLIYARGARGPROSERPHEARGALALEULEURGASPLEUASNGLYLEU 776  
 Db 2342 GTATGATCTGCGCGCGCCCTCTTCGAGCTATCTCTGAGAACTCAACGGGCTCAT 2401  
 QY 776 ETHRSERASPTYRGLIULEUUSERASPPROTHPRROGLIYLEPROSERPROARGSPGL 796  
 Db 2402 TACATCAGATTACAGAGCTCTCTCAGACCCCAACCTGGCATCCCGAGTCTCGAGATGA 2461  
 QY 796 ULUCYSVALALAGLYAGLNULEUTYRVALCYSGLINASPPROALALIEPHEGLINGLUAR 816  
 Db 2462 GCTCTGGG--GT-GGGCGCCGAGCTCTATGCTGCTGCGAGAGCCGCCCATATTCGAGGAGAG 2518  
 QY 816 GHISLEULSYRILLESERLEUGLYLYSGLYASNPHEGLYSERVALGILUEUCYSAR 836  
 Db 2519 ACACCTTAAGATACATCTTTGCTGGGCAAGGCAACTTGGCAGCGGAGCTGTGCGG 2578  
 QY 836 GTYASPPROLEUGLYASPSANTHRGLYPROLEUVALALAVALLYSGLINLEUGLINHIS 856  
 Db 2579 CTATGACCCCTCGGGGAGACATACGGAGCCCTGCTGCAATAACAGCTACAGCACAG 2638  
 QY 856 RVAL-PROASPGLINGLINARGASPHGELINARGGLUGLILEGILILEULEULYSALALEUHHIS 876  
 Db 2639 CGG-GCCAGACACAGAGGAGCTTCAGCGGAGATTCAGATCTTAAGGCTCTGCACA 2697

QY 876 ERAASPHELLEVALLYSTYRARGLYVALSERTYRGLYPROGLIYARGINSERLEUARGL 896  
 Db 2698 GCGACTTTCATGTCAGATACAGCGGAGATGACCTATGAGGCGCAGGTGCGCACAGCCCTGCGGT 2757  
 QY 896 EUVALMELGLUTYRLEUPROSERGLYCYSLERARGASPLEU-LEUGLINARGHIS---Ar 914  
 Db 2758 TGTGATGATGACTGCTGCCAGCGGCTCTCCGAGAG-CTTCCTGCACAGCGCATGCGCGG 2816  
 QY 914 GGLYLEUHHISTHRASPARGLEULEULEUPHEALATRPSGLINILECYLSYSGLYMETGLUTY 934  
 Db 2817 CGC-CTGCACACCGACCCCTACTGCTGTGCTGCGAGATCTTCGCAAGGCGCATGGAATA 2875  
 QY 934 RLEUGLIYALARGARGCYVALHISARGASPLEUVALALARGASNILEULEUVALGLIUSE 954  
 Db 2876 CTTGGGTGCGCGCGCTGCTGACACCGTACCTGGCTGCGGCAACATCTTGGTGGAGAG 2935  
 QY 954 RGLUALAHISVALLYSILEALIASPPHEGLYLEUALALYSLEULEUPROLEUGLYLSAS 974  
 Db 2936 CGAGGCTCATGTGAAGATCGCGGAGCTTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2995  
 QY 974 PTYTYRVALVALARGGLUPROGLYGLINSERPROLLEPHEPTYRVALAPROGLISERLE 994  
 Db 2996 CTACTACCTGCTGCGCGAGCTGCGCAAGCCCATCTTTGGATGCTGCGCGAGTCCCT 3055  
 QY 994 USERASPSANLLEPHESERARGINSERASPSVALTRPSPHEGLYVALLEUITYRGL 1014  
 Db 3056 ATCTGACACACTCTTCTCCGCCCAATGACGTGTGAGAGCTTGGAGAGTGTGTGACGA 3115  
 QY 1014 ULEUPHERTHYRCYASAPLYSERYSERPROSERALAGLUBHEUARGMETMEL 1034  
 Db 3116 GCTCTTCACCTACTGCGCAGACAGCTGACGCCCATCCGCTGAGTTCCTGCCACATGATGGG 3175  
 QY 1034 YPROGLUARGGLUGLYPROPOLEUCYARGLEULEUNGILUEUVALAGLUGLYARGAR 1054  
 Db 3176 GCCTGAGCGTAAAGAGACCCCGCTGCTGCGCTCTGAGAGTGGTGGAGAGGCGCGACG 3235  
 QY 1054 GLEUPROPROPROTHRYCSPROTHGLIUALGILINGLULEUWETGLINLEUCYSTYRPA 1074  
 Db 3236 CCTCCACACACTCCACACTCCGCCACGAGAGTTCAGAGACTCATGACAGCTGCTGCGC 3295  
 QY 1074 APRO 1075  
 Db 3296 GCCC 3299  
 RESULT 3  
 MUSJAK3H  
 LOCUS  
 DEFINITION Mus musculus (clone 32D5-1). protein tyrosine kinase (JAK3) mRNA.  
 ACCESSION L33768  
 VERSION L33768.1 GI:508533  
 KEYWORDS protein tyrosine kinase.  
 SOURCE Mus musculus cDNA to mRNA.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 4016)  
 AUTHORS Rane,S.G. and Reddy,E.P.  
 TITLE JAK3: a novel JAK kinase associated with terminal differentiation  
 of hematopoietic cells  
 JOURNAL Oncogene 9 (8), 2415-2423 (1994)  
 MEDLINE 9430920  
 PUBMED 7518579  
 FEATURES  
 source  
 location/Qualifiers  
 1..4016  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="32D5-1"  
 /cell\_line="32Dc13"  
 /tissue\_l1b="lambda gtl1"  
 7..3906  
 /note="protein tyrosine kinase"  
 /codon\_start=1

```

/product="JAK3"
/protein_id="AAA2145.1"
/db_xref="GI:537607"
/translation="MAPSEEPRLIQRSCLSSSEAGALHVLPPRGPPQRLSFS
EGDYLABEDLCVRAKACALIPVHSLFALATEDSCWFPSPHLEEDVDVOLVRL
REYFPMFGLETCNHRGLRKLDTSLDLHLVLEHLHAQRHSDLSVGRPLVGLSMKFG
EFLSLAVLDLAQARQAOQRPGLLKTVAQKALDPSLFDVDOGQNFVRRRLRRFV
LALRWSPARPATTAHQOVLDETLRLPAATETTFVGLPAQOEERGLVRRADNGIS
MSSGQDEVLGILRSVGRBAMRLVKRPVTRIKVSPRACCTGGGCVLKLKQPG
MELPGLTWGVTWRVRLDRGRGTEDSGMDPKIRKVPVPSPLSSPPPKRMWEG
GRRGAGSRSVPMILSLFLPFENGRQGFNSSGCGTHEVAPMPKQKMSACL
CLSSAVIKGRVPLRLKRLCSLFLQLEFQTFEDPFDVDSIKQPTCGSREHRLVYTRM
DGHLEAEFPGLEALSFVALVDGYRLCLDSNHRVCKEVAAPRLIEBAELCHGRT
LGASWQOQWGMWAARTVLGTMWLSMPRLDIAHKLKAASLPGYILRRSPDVAL
FLTACVQTPPLGPDYKGLIROPDSGAFSLVASASTEGSCDLEEGLEVDGAL
NLTSCCAPRKPKESNLIVVRGCTPAPRAGCSPPSCALFOLSFPHTPDSCHEMNG
GSPTFKIFGRRRREVVDGETHDESVLLKVDNRHNMESFLQASLSMQSVPIHVLV
HGVCMAGDSIMQDEYVLCALIMYLKRGHLVSAKMLQVTKLALVLYLEDKGLPH
GVNSAKVILLAREGGDGNPFTIKLSPGVSPVLSLEMTDRIPMVAPCLOEAOTLC
LEADKWGFATTWEEVSGPAHITSLERPAKLIKFEQDQPLAKWTELGLITOCMA
YDGRSPSRFALRLDLNGLITSDYELLQPLHLASYLEMSCAMPALCLPGRHTEER
HLKYLISLCKGNFSELCRDPLDNTGRLVAVKLOHSGPPOORPORETOILKALH
SPFIYKRGVSYGPGROSRLVMEYLPSCGLRDLORHNAALHTDRILLFNAOICKM
EYLGARCVHRLAARNILVSEAHYKILRECLRALILPLKRYTYVREPGSPITWVA
PESLSNIFSRSDVMSFQVLYELFTTCDKSCSFAELRMGPERGRPLRLLEL
LABGRLLPRLPPTCTPEVDELQVLQVAPSRITTGQPSAP"
misc_feature
2199..3024
       /note="JAK homology domain II (JH2)"
misc_feature
3076..3876
       /note="JAK homology domain I (JH1)"
BASE COUNT      742 a 1203 c 1235 g      836 t
ORIGIN
      misc_feature
      180.00      Length:      4016
      Score:      7.77e-165      Matches:      529
      Percent Similarity:      97.24%      Conservative:      9
      Best Local Similarity:      97.24%      Mismatches:      9
      Query Match:      16.388      Indels:      15
      DB:      10      Gaps:      0
US-09-397-967-16 (1-1099) x MUSJAK3H (1-4016)
QY      535      ARGATGTAValValAspClyGluThrHisAspSerGlyValLeuLeuLeuValMetAsp      554
      |||||
DB      2257      AGCGGGAGAGTGTGATGTCAGACATGACTCGGAAGTCTCTCGTAAGGTCAATGGAC      2316
QY      555      SerATGHisATGAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnVal      574
      |||||
DB      2317      TCCAGACATCGGAACATGATGAGTCTTTTGGAAAGCGGCAAGCTTGATGAGCCACAGA      2376
QY      575      SerTYRrHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetVal      594
      |||||
DB      2377      TCCTACCCCGCAGCTGGTGTACTGACAGCGCTCTGTCATGGCTGGAGACATCATGGTG      2436
QY      595      GInGluPheValTYrLeuGluAlaAlaIleAspMetTYrLeuArgLYsArgGlyHisIleVal      614
      |||||
DB      2437      CAGGAATTTGTGTATCTAGAGACATTCATGATCTGGCGCAACCGTGGCCACCTGGTG      2496
QY      615      SerAlaSerTYrPlyLeuGluGlnValThrLYsGlnLeuAlaTYrAlaLeuAsnTYrLeuGlu      634
      |||||
DB      2497      TCAGCCAGCTGGAACTGACAGTCAAGAGCTGGCATATAGCCCTTAACACTTGGAG      2556
QY      635      AspLYsGlyLeuPheHisGlyAsnValSerAlaAlaGlyValLeuLeuAlaArgGlyGly      654
      |||||
DB      2557      GACCAAGGCTTCTCTACAGCGCAAGCTCTACAGCAGGAAGTGGCTCTCGCTGTGAGGGG      2616
QY      655      GLYAspGlyAsnProPheIleLYsLeuSerAspProGlyValIleSerProThrHisIleVal      674
      |||||
DB      2617      GGTGATGGGAATCCACCTTTCATTAAGCTGATGATCTGGGTGATGATCCACATGTGCTG      2676
QY      675      SerIleGluMetLeuThrAspArgIleProTYrValAlaProGlyCysLeuGlnGluAla      694
      |||||

```

```

DB      2677      AGCTGTGAATGCTCAACCGACAGATATACCTGGGTGGCCCCCGAATGTCTCCAGAGGCT      2736
QY      695      GInThrLeuCysLeuGluAlaAspLYsTYrGlyPheGlyAlaThrThrTrpGluValPhe      714
      |||||
DB      2737      CAGACACTGTGCTGGAGGCTGACAGTGGGGCTTTGGAGCCACACGCTGGAGAGGTGT -      2795
QY      715      GInArg -GlyProAlaHisIleThrSerLeuGluProAlaLYsLYsLeuLYsPheTYrG1      734
      |||||
DB      2796      CAGCGGGGAGCCCGCCACATCACCTCCGTGGAGCGCCCAAAAGCTGAATTCATGAC      2855
QY      734      uAspGlnGlyGlnLeuProAlaLeuLYsTrpThrGlyLeuAlaGlyLeuIleThrGlnCly      754
      |||||
DB      2856      GGACCGAGGACAGAGCTCCCGCTCTCAATGGACAGACTGGGGGAGCTTATACACAGAG      2915
QY      754      SheAlaTYrAspProGlyLYrArgArgProSerPheArgAlaIleLeuArgAspLeuAsnG1      774
      |||||
DB      2916      CATGGGCTATGATCTCGGCGCGGCGCCCTCTCCAGAGTATCTCAGAGACTTCACAGG      2975
QY      774      yLeuIleThrSerAspTYrGlyGlyLeuLeuSerAspProThrProGlyIleProSerProAr      794
      |||||
DB      2976      CCTCATTTACATAGATTTACGAGCTCTCCCA -GACCCACACCTGGCATCCGAGTCCCG      3034
QY      794      gAspGlyLeuCysValAlaGlyAlaGlnLeuTYrAlaCysGlnAspProAlaIlePheG1      814
      |||||
DB      3035      AGATGAGCTGTGCG -T -GGCGCCAGCTGATGCTCTCCAGAGACCCCGCATATTC -GA      3089
QY      814      uGluArgHisLeuLYsTYrIleSerLeuLeuGlyLYsGlyAsnPheGlySerValGluLe      834
      |||||
DB      3090      GAGAGACACCTTAAATGATCATCTCTTGTGGCGAAGGCAACCTTGGACAGCTGGAGACT      3149
QY      834      uCysArgTYrAspProLeuGlnLYsAsnThrGlyProLeuValAlaValLYsGlnLeuG1      854
      |||||
DB      3150      GTGCGGCTATGACCCCC -TG -GACATATACGGAGACCTCGGTGGAGAGTAACAGCTACA      3206
QY      854      nHisSerVal -ProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLYsAla1      874
      |||||
DB      3207      GCACAGCGG -GCCACACGACGAGAGAGGACTTCCAGCGGAGATTCAGATCCCTTAAGGCTC      3265
QY      874      eunHisSerAspPheIleValLYsTYrArgGlyValIleSerTYrGlyProGlyArgGlnSer1      894
      |||||
DB      3266      TGCACAGGAGCTTCAATCTCAAGTACCGGGAGAGTCAAGTATGGGCGAGGTGCCAGAGCC      3325
QY      894      euATGLeuValMetGluTYrLeuProSerGlyCysLeuArgAspLeu -LeuGlnArgHis      913
      |||||
DB      3326      TCGGCTTGATGATGAGTACCTGCCACAGCGCTGCGCTGCGAGA -CTTCCTGACAGCGCAT      3384
QY      914      - -ArgGly -LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLYsGlyMet      932
      |||||
DB      3385      CGCGGGGCTGACACACCGGACCGCTCTACTGCTGTTCCTGGCAGATCTGCAGAGGCAATG      3444
QY      933      GInTYrLeuGlnGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuVal      952
      |||||
DB      3445      GAGTACCTGGGTGGCGCGCGCTGCGCTACCGGTAACCTGGCTGGCGCCACATCTTGCTG      3504
QY      953      GluSerGlnAlaHisValLYsIleAlaAspPheGlyLeuAlaLYsLeuLeuProLeuG1Y      972
      |||||
DB      3505      GAGAGCGAGGCTCATGTGAAGATCGCGGACTTCGGCTCGCTAAGCTGAGTGGCCCTGGGA      3564
QY      973      LYsAspTYrTYrValValArgGlyLupProGlyGlnSerProIlePheTYrTYrAlaProGlu      992
      |||||
DB      3565      AAGGACTACTACGTGGTCCCGGAGCGGCGCCAAAGCCCAATCTTTTGTGTATGCCCGCGAG      3624
QY      993      SerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeu      1012
      |||||
DB      3625      TCCTATCTGACACATCTTCTCCGCGCATCTACCTGTGAGACTTCGAGAGTGGTGTG      3684
QY      1013      TYrGlyLeuPheThrTYrTYrCysAspLYsSerCysSerProSerAlaGluPheLeuArgMet      1032
      |||||
DB      3685      TACAGGCTTTCACCTTACTCGAGACAAAGGCTGCGCCATCCGCTGAGTTCCTGCGCATG      3744
QY      1033      MetGlyProGluArgGlyGlyProProLeuCysArgLeuLeuGlnLeuAlaGlyGly      1052
      |||||
DB      3745      ATGGGCGCTGAGGCTGAGAGACCCCGCTCTGCGGCTCTGGAGCTGTGGGAGAGGCG      3804

```



[illegible]

Alignment Scores:		9.01e-53	Length:	3778
Pred. No.:	65.00	Matches:	99	
Score:	98.02%	Conservative:	0	
Percent Similarity:	98.02%	Mismatches:	1	
Best Local Similarity:	98.02%	Indels:	2	
Query Match:	5.91%	Gaps:	0	
DB:	10			
US-09-397-967-16 (1-1099) x RATJAK3 (1-3778).				
QY	919	AspargLeuLeuPheAlaTrpGlnIleCysIysGlyMetGluTyrLeuGlyAlaArg	938	
DB	2818	GACCCCTCAAGCTCTTGGCCCTGGAGATCTGCAAGGCGCATGGAATACCTGGCGCACGC	2877	
QY	939	ArgCysValIHisArgAspLeuAlaIalaArgAsnIleLeuValGluSerGluAlaHisVal	958	
DB	2878	CCCTCGTCGTCATCTGACCTGGCTGGCGGCAACATCTCGTGGAGACATGAGCGCATGTG	2937	
QY	959	LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal	978	
DB	2938	AAGATTGCTGACTTGGCGCTCGCTAAGCTGGCCCTCTTGGAAAGACTACTACTGGTC	2997	
QY	979	ArgGluProGlyIysIleSerProIlePheTrpTyrIleAlaProGluSerLeuSerAspAsnI	998	
DB	2998	CGCGT-GCCGGGSCCAAGTCCCATCTTCTGGTACGCCCTTGATGCCCTATCTGACACAT	3056	
QY	998	ePheSerArgGlnSerAspValTyrPheSerPheGlyValValLeuTyrGluLeuPheThrTy	1018	
DB	3057	CTTCTGCCCAATCTGACGTGTGAGAGTTCGGCGTGGCTGTGATGACCTCTTACACCTA	3116	
QY	1018	r 1018		
DB	3117	C 3117		
RESULT 5				
LOCUS	AX203087	778 bp	DNA	linear
DEFINITION	Sequence 8 from Patent WO0152892.			PAT 30-AUG-2001
ACCESSION	AX203087			
VERSION	AX203087.1	GI:15392445		
KEYWORDS				
SOURCE	human.			
ORGANISM	human sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
TITLE	1 (bases 1 to 778)			
JOURNAL	Jak/stat pathway inhibitors and the uses thereof			
	Patent: WO 0152892-A 8 26-JUL-2001;			
	GENZYME CORPORATION (US)			
FEATURES	Location/Qualifiers			
source	1..778			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
BASE COUNT	131 a 265 c 219 g 163 t			
ORIGIN				
Alignment Scores:				
Pred. No.:	2.14e-50	Length:	778	
Score:	62.00	Matches:	96	
Percent Similarity:	97.96%	Conservative:	0	
Best Local Similarity:	97.96%	Mismatches:	2	
Query Match:	5.64%	Indels:	2	
DB:	6	Gaps:	0	
US-09-397-967-16 (1-1099) x AX203087 (1-778)				
QY	938	ArgArgCysValIHisArgAspLeuAlaIalaArgAsnIleLeuValGluSerGluAlaHis	957	
DB	103	CGCCGCTGCTGACCGCGACCTGGCGCGGCAACATCTCTCGTGGAGACGAGGCGCAC	162	
QY	958	ValIysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVal	977	



Accession	Gene	Protein	Length	Score	Similarity	Local Similarity	Query Match	DB
Db 163	GRCAAGATCGCTGACTTGGCCCTAGCTAAGTCTGCTGCCGCTTGAC-AAAGCTACTAGCT	221						
Qy 977	1VALARGLUProGlyInSerProIlePheTrpTyrAlaProGluSerLeuSerAspAs	997						
Db 222	GCTCGCGACGACGCGCCAGAGCCCATTTCTTGATATGCCCGAATCCCTCTCGACAA	281						
Qy 997	11LEPHESEIRGINSERASPAVALTPSERPHEGLYVALLEUTYRGLULEUPHERH	1017						
Db 282	CATCTTCTCTGCGCCAGTACAGCTCTGTGAGCTTGAGGCTGTCTGTACGAGCTCTTCAC	341						
Qy 1017	RTYRCYASPLYSERCYSSERPROSERALAGLUPEHLEUARGMETMETGLY	1034						
Db 342	CTACTGCGACAAAGCTGACAGCCCTCGGCCGAGTTCCTCGGATGATGGGA	393						
RESULT 6	AX203084	2914 bp	DNA	linear	PAT 30-AUG-2001			
LOCUS	Sequence 5 from Patent WO0152892.							
DEFINITION	AX203084							
VERSION	AX203084.1	GI:15392441						
KEYWORDS								
SOURCE	human.							
ORGANISM	Homo sapiens							
REFERENCE	Makarewicz; Metazoa: Chordata; Craniata; Vertebrata: Euteleostomi;							
AUTHORS	1 (bases 1 to 2914)							
TITLE	Jak/stat pathway inhibitors and the uses thereof							
JOURNAL	Patent: WO 0152892-A 5 26-JUL-2001;							
FEATURES	GENZYME CORPORATION (US)							
source	Location/Qualifiers							
	1..2914							
	/organism="Homo sapiens"							
	/db_xref="taxon:9606"							
BASE COUNT	566 a 907 c 809 g 632 t							
ORIGIN								
Alignment Scores:								
Pred. No.:	6.15e-50	Length:	2914					
Score:	62.00	Matches:	96					
Percent Similarity:	97.96%	Conservative:	1					
Best Local Similarity:	97.96%	Mismatches:	1					
Query Match:	5.64%	Indels:	0					
DB:	6	Gaps:	2					
US-09-397-967-16 (1-1099) x AX203084 (1-2914)								
Qy 938	ARGARGCYSAVALHISARGASPLEUALALARGASNIILEUVALGUSERGLUALAHIS	957						
Db 1772	CCCCCTGGCTGCACCGCGACCTGGCCGCCGGAACATCCTCTGTGAGAGGAGGACACAC	1831						
Qy 958	VALYSILIEALIAAPPHGCLYLEUALALYSLEULEUPRO-LEUGLYLSASPTYRTRYVA	977						
Db 1832	GTCAAGATGCTGACTTGCGCTAGCTAAGCTGCTGCCGCTTGAC-AAAGACTACTAGCT	1890						
Qy 977	1VALARGLUProGlyInSerProIlePheTrpTyrAlaProGluSerLeuSerAspAs	997						
Db 1891	GGTCCGCGACGACGAGCCAGAGCCCATTTCTGTGTATGCCCGCAATCCCTCTGAGCAA	1950						
Qy 997	11LEPHESEIRGINSERASPAVALTPSERPHEGLYVALLEUTYRGLULEUPHERH	1017						
Db 1951	CATCTTCTCTGCGCCAGTACAGCTCTGTGAGCTTGAGGCTGTCTGTACGAGCTCTTCAC	2010						
Qy 1017	RTYRCYASPLYSERCYSSERPROSERALAGLUPEHLEUARGMETMETGLY	1034						
Db 2011	CTACTGCGACAAAGCTGACAGCCCTCGGCCGAGTTCCTCGGATGATGGGA	2062						
RESULT 7	HSU31601	3620 bp	mrna	linear	PRI 24-OCT-1995			
LOCUS	Human tyrosine protein kinase (JAK3B) splice variant mRNA, complete							
DEFINITION	cs.							
ACCESSION	U31601							

VERSION	KEYWORDS	UT31601.1	GI:1039418
SOURCE	ORGANISM	Homo sapiens.	
REFERENCE	AUTHORS	Liu, E.T., A kinase-deficient splice variant of the human JAK3 is expressed in hematopoietic and epithelial cancer cells J. Biol. Chem. 270 (42), 25028-25036 (1995)	
TITLE		2 (sites)	
JOURNAL		Wittuhn, B.A., Silvennoinen, O., Miura, O., Lai, K.S., Cwik, C., Liu, E.T., and Ihle, J.N.	
MEDLINE		Involve ment of the Jak-3 Janus kinase in signalling by Interleukin 2 and 4 in lymphoid and myeloid cells	
PUBMED		Nature 370 (6485), 153-157 (1994)	
REFERENCE		3 (bases 1 to 3620)	
AUTHORS		Liu, E.T.	
TITLE		Direct Submission	
JOURNAL		Submitted (14-JUL-1995) Edison T. Liu, Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599-7295 USA	
MEDLINE		Wittuhn, B.A. Nature 370, 153-157, 1994.	
PUBMED		Location/Qualifiers	
REFERENCE		1. .3620	
AUTHORS		/organism="Homo sapiens"	
TITLE		/db_xref="taxon:9606"	
JOURNAL		/chromosome="19"	
MEDLINE		/map="19p12"	
PUBMED		/clone="SKB-3"	
REFERENCE		/cell_type="SKB-3 breast cancer cells"	
AUTHORS		1. .3620	
TITLE		/gene="jak3b"	
JOURNAL		96. .3380	
MEDLINE		/function="tyrosine protein kinase"	
PUBMED		/codon_start=1	
REFERENCE		/product="jak3b"	
AUTHORS		/protein_id="AAC50226.1"	
TITLE		/db_xref="gi:1039419"	
JOURNAL		/translation="MAPSEETPLIPQSCSLSTEGALHVLIPARGPPEPRLSFSFGDHLAEGLGVOAKKAGSLGIVHSLPALNEDTSCPPSHSFSDASTOVLKTRIREFPNMNGLEKCHRRGKRLKLAIIIDLPYLEHLFQHNHSDVSGRLPGVLSLKDGECLSIAVLDLARMADEQAQRGELKTVSTKACIPSLRLDGLSVTRRRIRTVRRAIRVAACQADNRSLMAKYMIDLERIDPAGAEETPHVGLGALGDHGLGLRVAGDGGIATVGEDEVLQPCDFEPIVDISTKOAPRQAGEHRLVYTRTDNOILAEFEGLEALSFVALVDGFRLLTDSOHFFCEKVAAPRLLEEAGVOCGPIITDPAINKKSGSRGSAVLRSPQDFPSFLTVGVONLPDVKGLIRSPGCTFLADGSLNSPHSLRELATCMWDGLHVDGAVALTSSCIPRPEKKNLIYVQNGHSPPTSLVQPOSVOL	
MEDLINE		SOMPFHPIPADSLHEWHENLGHSPTKIRRCRHEVDGEARKTEVLLKWDAAHKNCMLQVYKQALAVYLEDGALPHGNVSAKRVLLAEEGADGSPFIKLSDPGSPAVLSLEMDRITVVAPECLREKQVTLSEADKKFGCATWEYFSGVTMPISALDPKLOFEDRQQLPAKRWELALILQOCMAVEPVQSPRAVIRIOLNLSISDELLSDPILGALAPRDGAGNAGLYACODPITFEERHLKITSOLGKNGSVELCAYDPDGLDMTCAVLAVKLDHSGPQDQDFOREQIILKALSDPILVKRGVSPGQSLRLWNEYLPSCGLRDLQOHRRARLDASRLLSOIOCKGMEYLSGRCRVRLAARLLISSEBAHKVIADFGKALLPIDPKDYVVRREGOSPIITVYAPESLSDNITSROSDWVSFGVILYELPTVYDSCSPSAETLRMGCEKRVYFALCRLELLEEGQRILPAPPAQPAVSAAGLASVQSOSVDMAGVSGKPAQA"	
REFERENCE		BASE COUNT	670 a 1152 c 1076 g 722 t
AUTHORS		ORIGIN	
TITLE		Alignment Scores:	
JOURNAL		Pred. NO.:	7.31e-50
MEDLINE		Score:	62.00
PUBMED		Percent Similarity:	97.96%
REFERENCE			Length: 3620
AUTHORS			Matches: 96
TITLE			Conservative: 0

Best Local Similarity: 97.96% Mismatches: 1  
 Query Match: 5.64% Indels: 2  
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x HSU31601 (1-3620)

OY 938 ArgatgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957  
 |||||||  
 DB 2922 CGCCGCTCGGACCGGACCTGGCCGCCGGAACATCCCTCGGAGAGGAGGACACAC 2981  
 |||||||  
 OY 958 ValIysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977  
 |||||||  
 DB 2982 GTCAGATCGGCTGAGCTTGGCTTACGCTGCTGCGCTTGAC-AAAGACTACTAGT 3040  
 |||||||  
 OY 977 lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs 997  
 |||||||  
 DB 3041 GGTCCGCGAGCCGAGCCAGGCCATTTTCTGATGCCCCCGAATCCCTCGGACAA 3100  
 |||||||  
 OY 997 nIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheTh 1017  
 |||||||  
 DB 3101 CATCTTCTCGCCAGTACAGCTGTGAGAGCTTCGGGTCCTCTGACGAGACTTTCAC 3160  
 |||||||  
 OY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetGly 1034  
 |||||||  
 DB 3161 CTACTGCGACAAAGCTGCGCCCTCGGCGAGTTCTCGCGATGATGGGA 3212  
 |||||||

RESULT 8  
 AX203081 4064 bp DNA linear PAT 30-AUG-2001  
 LOCUS Sequence 2 from Patent WO0152892.  
 DEFINITION AX203081  
 ACCESSION AX203081  
 VERSION AX203081.1 GI:15392438  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 4064)  
 AUTHORS Vassios, G.  
 JOURNAL JAK/stat pathway inhibitors and the uses thereof  
 PATENT: WO 0152892-A 2 26-JUL-2001;  
 GENZYME CORPORATION (US)  
 FEATURES  
 source 1. 4064  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 746 a 1292 c 1168 g 858 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 8.02e-50 Length: 4064  
 Score: 62.00 Matches: 96  
 Percent Similarity: 97.96% Conservative: 0  
 Best Local Similarity: 97.96% Mismatches: 1  
 Query Match: 5.64% Indels: 2  
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AX203081 (1-4064)

OY 938 ArgatgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957  
 |||||||  
 DB 2922 CGCCGCTCGGACCGGACCTGGCCGCCGGAACATCCCTCGGAGAGGAGGACACAC 2981  
 |||||||  
 OY 958 ValIysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977  
 |||||||  
 DB 2982 GTCAGATCGGCTGAGCTTGGCTTACGCTGCTGCGCTTGAC-AAAGACTACTAGT 3040  
 |||||||  
 OY 977 lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs 997  
 |||||||  
 DB 3041 GGTCCGCGAGCCGAGCCAGGCCATTTTCTGATGCCCCCGAATCCCTCGGACAA 3100  
 |||||||  
 OY 997 nIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheTh 1017  
 |||||||

DB 3101 CATCTTCTCGCCAGTACAGCTTGAGAGCTTGGGCTGCTCTGTACAGACTCTTTCAC 3160  
 |||||||  
 OY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetGly 1034  
 |||||||  
 DB 3161 CTACTGCGACAAAGCTGCGCCCTCGGCGAGTTCTCGCGATGATGGGA 3212  
 |||||||

RESULT 9  
 HSU09607 4064 bp mRNA linear PRI 09-JUL-1994  
 LOCUS Human JAK family protein tyrosine kinase (JAK3) mRNA, complete cds.  
 DEFINITION Human JAK family protein tyrosine kinase (JAK3) mRNA, complete cds.  
 ACCESSION U09607  
 VERSION U09607.1 GI:508730  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1124)  
 AUTHORS Kawamura, M., McVicar, D.W., Johnston, J.A., Blake, T.B., Chen, Y.,  
 Lal, B.K., Lloyd, A.R., Kelvin, D.J., Staples, J.E., Ortaldo, J.R. and  
 O'Shea, J.  
 TITLE Molecular cloning of L-JAK, a Janus family protein-tyrosine kinase  
 expressed in natural killer cells and activated leukocytes  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (14), 6374-6378 (1994)  
 MEDLINE 94294384  
 PUBMED 8022790  
 REFERENCE 2 (bases 1 to 4064)  
 AUTHORS O'Shea, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-MAY-1994) John O'Shea, Leukocyte Cell Biology Section,  
 Fderick, MD 21702, USA  
 FEATURES  
 source 1. 4064  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="YT, HUT-78"  
 /cell\_type="natural killer cell, PMA activated T cells"  
 1. 4064  
 /gene="JAK3"  
 96.3470  
 /gene="JAK3"  
 /codon\_start=1  
 /product="JAK family protein tyrosine kinase"  
 /protein\_id="AA19626.1"  
 /db\_xref="GI:508731"  
 /translation="MAPPSEPTLIPORSCSLSTFAGALHVLPPARAPGPPORLSFS  
 FGDHLADLVCYQAAKASGILPYHSLEALTEDELSCWEPSPHIFSEVDASTOVLARI  
 RYFPNMFGLKCRHGRGLRKDLASALIDLPVLEHFAOHRSDLSVGRSLVGLSLKOG  
 ECLSLAVLDLARMEQAOARGELKTIVSKAKCLPSLRDLTGSLFVTRAIRTVR  
 RALPRVACQADRLSLMAKYIMDLERIDPAGAEIVHGVICPGLGHDGILGRVAG  
 GGIAMTGOGEVLQPCDFEIVDISIKQAPRGPAEHRVLVTRIDNOILEAEPG  
 LPEALSFALVDGFEFLTDSQHFCEKVPAPRLLEVAAGCHGPTLIDALINKLTG  
 GSRGSVYAIRSPQDPSPLITVCVONKPLADYKGLIRSPGTLVLVGLSPHSL  
 RELLATCMDSGLHYDVAVTLTSCIPRPKESNLIVQGHSPPTSLVQPOSOYL  
 SQMTFHKIPADSLKHEHNGHSTKTYRCGRHEVYVGEARKTEVLKVDAAHKNM  
 ESFLFAASLMSQVSRYHLVLHGVCMACDSTVMOEPVHLGALIMYLRKRGHGLVLSLE  
 LOVYQOLAVNLNLEADKGLPHGNVSAKVLILARGAGSPPIKLSPPGSPVLSLE  
 MLTRIPVWAECELEAQTSLLEADKGFQATVWEVSQVMTPLISALDPKKILOTFED  
 RQOLPAPKTEALILIOCMAYEPVORSEFAVIRDNLSLISDYELLSPDTGALVQ  
 KDGLNGQALVACODPTFEERHKLISQKGNFSGVEICRYDPLNHNALVAVNQ  
 LQHSQEDQORPQREIOTLKLKGLHSDFTVYKGVSGYGGREPLRLVMEYLSGCLRFL  
 OHRARLIDASRLILYSQICGMELSGRKYVNDLARNLTVSEAHVYIADGLAK  
 LPLDKDYVAVREPOSPITFWADESLSDNIFSQSDVMSGQVLYELFLFYCDKSCSP  
 SAELRMKGCEDVAPLCRLLELEEGRLAPAPCAPREHMLKLMCLMAFSPQDRPFS  
 SALGPDLMWNGSRGCTHAFTAHPEKHLISLFS"

BASE COUNT 746 a 1292 c 1168 g 858 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 8.02e-50 Length: 4064  
 Score: 62.00 Matches: 96  
 Percent Similarity: 97.96% Conservative: 0

FEATURES	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	FUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	
1	13562	Human JAK3 gene, complete cds.	U70065	1	GI:1800224	Homo sapiens.	Homo sapiens.	1	Boese, A., O'Shea, J.J., and Johnston, J.A.	Genomic sequence, organization, and chromosomal localization of human JAK3	Genomics 37 (1), 57-61 (1996)	97079659	8921370	2	Riedy, M.C., Dutra, A.S., Blake, T.B., Modi, W., Lal, B.K., Davis, J., Boese, A., O'Shea, J.J., and Johnston, J.A.	Direct Submission	Submitted (09-SEP-1996)	NIAMS, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA	1. 13562

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.1"
/gene="JAK3"
join(201..384,719..842,1141..1255,1462..1607,2025..2317,
2865..2987,3082..3350,4264..4450,6303..6430,6630..6761,
7418..7503,7918..8044,8808..8939,9019..9171,9302..9452,
10839..110978,11059..11256,12009..12125,12400..12969)
/gene="JAK3"
/codon_start=1
/product="JAK3"
/protein_id="AAC50950.1"
/db_xref="GI:1800225"
/translation="MAPSEEPPLIPORCSCLSTPEAGALHYLLPARGPQPRLSFS
SCDHLADLCVOAKAGTLPVYHSIFALATELSCMPSPSHFSVEDASTOVLRYI
RSFYPMWFGLEKCHRFGLRKDLASNILDLPIYEHFAQHRSDIVSGRLPVGLSKEO
GECLSLAVLDLARMAREQAKRGELIKTYSKACLPSPSLDLIOGLSPYTRAIRTV
RRALPRVACQADHRSHIMAKYIMDLERLPAGAAETFHVGLFGALGHHGLLFRAG
DGIAMTQGEVILQPCDPEIIVDISIKQAPRVGGEHRLVTVRTDQILLEPFG
LPEALSFVALVDGYFRITDSDHFCFEKQVAPRLEBEVAEQHGPIITLDFAINKLKG
GSRGSLVLRSPQDPSFLIVCVONPLGPDYKGLIRSPGTGFLVLGSPHSL
RELLATCMQDGLHDGVAATLTSCIPRPEKSNLIVORGHSPPTSSIVQSOYOL
SQMTPKRPADSLFMHENVLGHGSETKIYRGCRHEVVDGEARKTEVLTKRKGHLVPASMK
ESFLKASLSQVSYRHLVYLHGVCMAGCSTWQEPVHIGAITMYLRKRGHLVPASMK
LQVVKQLAVLNLNLEDKGLPHGNVSAKUYLLAREGADSGSPPTKLSDPGVSPAVLSLE
MLNDRIPWAVECLREAOITLSEADKMGFATWEVFSGYTMPISALDPAKKIQPED
RQOLPAKMTLALLIQOCMAVEPVQRPSPRAVIRDLNLSISDYELSLSPITFGALAP
RDGLMNGAOLYACODPTIFEEHRLKYISQLGKNGFSGVELCRYPDLGDNGLAVAKQ
LQHGSPQORDFOREIOILKALHSDFIYKRGVSYGPELRLVMEYLPSCGLRFL
ORHARRIDASRLILYSOICGMGYLSRRCVRDLAARNILVESEAHVYIADFGLAK
LLPLDKIYYVREPGQSPRFVYAPESQISDNIFRSQDPSVFGVLLKELFYCDKSCSP
SAELPRLMGCERDVPALCRLLLELEBQRIAPRPACAEVHELMKLCMAWSPQDRSE
SALGPOLDMLMSGSGCTHAFTAHPRGKHHSFS"
201..384
/gene="JAK3"
/number=1
385..718
/gene="JAK3"
/number=1
719..842
/gene="JAK3"
/number=2
843..1140
/gene="JAK3"
/number=2
1141..1255
/gene="JAK3"
/number=3
1256..1461
/gene="JAK3"
/number=3
1462..1607
/gene="JAK3"
/number=4
1608..2024
/gene="JAK3"
/number=4
2025..2317
/gene="JAK3"
/number=5
2318..2864
/gene="JAK3"
/number=5
2865..2987
/gene="JAK3"
/number=6
2988..3081
/gene="JAK3"
/number=6
3082..3350
/gene="JAK3"

```

```

/number=7
3351..4263
/gene="JAK3"
/number=7
4264..4450
/gene="JAK3"
/number=8
4451..6302
/gene="JAK3"
/number=8
6303..6430
/gene="JAK3"
/number=9
6431..6629
/gene="JAK3"
/number=9
6630..6761
/gene="JAK3"
/number=10
6762..7417
/gene="JAK3"
/number=10
7418..7503
/gene="JAK3"
/number=11
7504..7917
/gene="JAK3"
/number=11
7918..8044
/gene="JAK3"
/number=12
8045..8807
/gene="JAK3"
/number=12
8808..8939
/gene="JAK3"
/number=13
8940..9018
/gene="JAK3"
/number=13
9019..9171
/gene="JAK3"
/number=14
9172..9301
/gene="JAK3"
/number=14
9302..9452
/gene="JAK3"
/number=15
9453..10839
/gene="JAK3"
/number=15
10839..10978
/gene="JAK3"
/number=16
10979..11058
/gene="JAK3"
/number=16
11059..11256
/gene="JAK3"
/number=17
11257..12008
/gene="JAK3"
/number=17
12009..12125
/gene="JAK3"
/number=18
12126..12399
/gene="JAK3"
/number=18
12400..13562
/gene="JAK3"
/number=19

```



AC073750 223734 bp DNA linear HTG 29-JUN-2000  
LOCUS Mus musculus clone RP23-330D8, WORKING DRAFT SEQUENCE, 28 unordered  
DEFINITION pieces.  
AC073750  
AC073750.1 GI:8810367  
VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1. (bases 1 to 223734)  
DOE Joint Genome Institute.  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 1863672  
Center clone name: RP23-330D8

-----  
Summary Statistics  
Consensus quality: 196566 bases at least Q40  
Consensus quality: 208241 bases at least Q30  
Consensus quality: 210965 bases at least Q20  
Estimated insert size: 208000: agarose-fp estimation  
Estimated insert size: 221034: sum-of-contigs estimation  
Quality coverage: 8.01 in Q20 bases; agarose-fp estimation  
Quality coverage: 7.54 in Q20 bases; sum-of-contigs estimation.  
NOTE: This is a 'working draft' sequence. It currently  
\* consists of 28 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1027: contig of 1027 bp in length  
1127: gap of unknown length  
1128 1127: contig of 1034 bp in length  
2162 2161: gap of unknown length  
2262 2261: gap of unknown length  
3709: contig of 1448 bp in length  
3710 3809: gap of unknown length  
3810 4864: contig of 1055 bp in length  
4865 4964: gap of unknown length  
5985: contig of 1021 bp in length  
5986 6085: gap of unknown length  
6086 7691: contig of 1606 bp in length  
7692 7791: gap of unknown length  
7792 10317: contig of 2526 bp in length  
10318 10417: gap of unknown length  
10418 11756: contig of 1339 bp in length  
11757 11856: gap of unknown length  
11857 13223: contig of 1367 bp in length  
13224 13323: gap of unknown length  
13324 16114: contig of 2791 bp in length  
16115 16214: gap of unknown length  
16215 20778: contig of 4564 bp in length  
20779 20878: gap of unknown length  
20879 23553: contig of 2675 bp in length  
23554 23653: gap of unknown length  
23654 26667: contig of 3014 bp in length  
26668 30222: gap of unknown length  
30222 30322: contig of 3455 bp in length  
30322 gap of unknown length

\* 30323 34899: contig of 4577 bp in length  
\* 34800 34989: gap of unknown length  
\* 35000 39609: contig of 4610 bp in length  
\* 39610 39709: gap of unknown length  
\* 39710 46418: contig of 6709 bp in length  
\* 46419 46518: gap of unknown length  
\* 46519 50029: contig of 3511 bp in length  
\* 50030 50129: gap of unknown length  
\* 50130 56982: contig of 6853 bp in length  
\* 56983 57082: gap of unknown length  
\* 57083 66435: contig of 9353 bp in length  
\* 66436 66535: gap of unknown length  
\* 66536 73174: contig of 6639 bp in length  
\* 73175 73274: gap of unknown length  
\* 73275 87407: contig of 14133 bp in length  
\* 87408 87507: gap of unknown length  
\* 87508 100238: contig of 12731 bp in length  
\* 100239 100338: gap of unknown length  
\* 100339 119196: contig of 18858 bp in length  
\* 119197 119296: gap of unknown length  
\* 119297 140119: contig of 20823 bp in length  
\* 140120 140219: gap of unknown length  
\* 140220 169227: contig of 29008 bp in length  
\* 169228 169327: gap of unknown length  
\* 169328 198979: contig of 29652 bp in length  
\* 198980 199079: gap of unknown length  
\* 199080 223734: contig of 24655 bp in length.  
Location/Qualifiers  
1. 223734  
source  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-330D8"  
/clone\_1b="RP23 mouse BAC library 23"  
BASE COUNT 57507 a 53185 c 54134 g 56203 t 2705 others  
ORIGIN

Alignment Scores:  
Pred. NO.: 1.98e-48 Length: 223734  
Score: 62.00 Matches: 62  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.64% Indels: 0  
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x AC073750 (1-223734)

OY 1 MetaIaProserGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
|||||  
Db 124160 ATGGACCTCCAGAGAGACACACTGTGATCCCTGACGCTTGCAGCCTGCATCC 124219

OY 21 SerGIuaIaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40  
|||||  
Db 124220 TCAGAGCGAGGAGCCCTGCATGCTCTCTCCCGGAGCCCTCCCGGAGCCCTCCCGAGGA 124279

OY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGlnAspLeuCysValArgAlaAlaAlaAla 60  
|||||  
Db 124280 TTGTATTCTCTTTGGGACTACTTGCTGCTGAGATTTATGTGTGCCAGCTGCCAAGGCC 124339

OY 61 CysGly 62  
|||||  
Db 124340 TGTGCT 124345

RESULT 14  
LOCUS AX203086 795 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 7 from Patent WO0152892.  
ACCESSION AX203086  
VERSION AX203086.1 GI:15392444  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.



[illegible]



[illegible]

LQVYKQLAYALNYLEDKGLPHGNVSAKRVLLAREGADGSPPTIKISDPGVSPAVLSLE  
 MLDRIIPWVAPECLREAOQLSLKADKWGATVWEFSGVMTPTISALDPAKLIQFTED  
 RQOLPAPKWTLEALLIQOCMAVEPVORPSPRAIRDLNLSISDYELLSDPPGALAP  
 RDGLMGAQIYACODPTIEERHLKYISOLGKNFSGVSELCRDYPLGDMTGAIVAKO  
 LQHSQDQORDFOREIOILKALHSDIYVGVSGPGQSLRLVMEYLPSCGLRPL  
 ORRRARLDASRLIYSSOICKMEYIGSVRCVRRDLAANILVESAHHKIDFGIAK  
 LPLDLDIYVYREPGOSPFIWYAPESLSNIPFSRODWSFGVYVTELELFTYCDKSGSP  
 SAEFLRMGCEKRVPAIACRLLELLEGQRLPAPACPAEVHLMKLCMAPSPQDRPSE  
 SAGPQLDMIMSSRGCEHAFVAPHEGRHNSLFS"

## variation

2579 /gene="JAK3"  
 /frequency="0.36"  
 /replace="t"

## variation

2597 /gene="JAK3"  
 /frequency="0.43"  
 /replace="a"

## variation

2599 /gene="JAK3"  
 /frequency="0.37"  
 /replace="a"  
 2619. .2778  
 /rpt\_family="L2"  
 /rpt\_type="dispersed"

## repeat\_region

## variation

2653 /gene="JAK3"  
 /frequency="0.32"  
 /replace="t"

## variation

2654 /gene="JAK3"  
 /frequency="0.04"  
 /replace="a"

## variation

3003 /gene="JAK3"  
 /frequency="0.07"  
 /replace="c"

## variation

3030 /gene="JAK3"  
 /frequency="0.02"  
 /replace="t"

## variation

3076 /gene="JAK3"  
 /frequency="0.07"  
 /replace="a"

## repeat\_region

3095. .3200  
 /rpt\_family="MIR"  
 /rpt\_type="dispersed"

## variation

3106 /gene="JAK3"  
 /frequency="0.06"  
 /replace="a"

## variation

3186 /gene="JAK3"  
 /frequency="0.06"  
 /replace="c"

## variation

3187 /gene="JAK3"  
 /frequency="0.06"  
 /replace="a"

## variation

3211 /gene="JAK3"  
 /frequency="0.06"  
 /replace="a"

## variation

3385 /gene="JAK3"  
 /frequency="0.06"  
 /replace="a"

## variation

3442 /gene="JAK3"  
 /frequency="0.04"  
 /replace="t"

## variation

## Alignment Scores:

Pred. No.:	7.89e-27	Length:	21323
Score:	39.00	Matches:	39
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.55%	Indels:	0
DB:	9	Gaps:	0

US-09-397-967-16 (1-1099) x AF513860 (1-21323)

Oy 989 TyraIaProgluSerIeuSerAspaSnIlePhSeSerArgGlnSerAspValTirpSerPhe 1008

Db 16179 TATGCCCCCGAATCCCTCGACAAACATCTCTCGCCAGTCACACGCTCGAGCTTC 16238

Oy 1009 GlyValValIeuIyrGluIeuPheThrIyrCysAspLysSerCysSerProSerAla 1027

Db 16239 GGGGTGCTCTGACGAGCTTCCTACCTACTGCGACAAAGCTGACGCCCTCGGCC 16295

## RESULT 18

AC007201/c 41006 bp DNA linear PRI 03-APR-1999

LOCUS Homo sapiens chromosome 19, cosmid R34383, complete sequence.

DEFINITION AC007201

VERSION AC007201.1 GI:4558769

KEYWORDS HTG.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE

AUTHORS

Lamerdin,J.E., McCready,P.M., Skowronski,E., Scott,D.,  
 Burkhardt-Schultz,R., Gordon,L., Dias,J., Sakladakis,G.,  
 Sillwagen,S., Phan,H., Velasco,N., Do,L., Regata,W., Terry,A.,  
 Dangnan,L., Erlar,A., Christensen,M., Georgescu,A., Avila,J.,  
 Attix,C., Andreise,T., Amico-Keller,G., Coefield,J., Duarte,S.,  
 Lucas,S., Bruce,R., Thomas,P., Quan,G., Krommiller,B., Arellano,A.,  
 Sanders,C., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S. and  
 Carrano,A.V.

Sequence analysis of a 5.7 Mb region in 19p13.1  
 Unpublished  
 2 (bases 1 to 41006)  
 Lamerdin,J.E.  
 Direct Submission  
 Submitted (03-APR-1999) Joint Genome Institute, Lawrence Livermore  
 National Laboratory, 7000 East Ave., Livermore, CA 94551, USA  
 Map and sequence oriented from p telomere to centromere. Cosmid  
 R34383 overlaps cosmid R19847 (AC005952) to the left from bases 1  
 to 2,923 of this accession, and overlaps cosmid R31408 (AC005796)  
 to the right from bases 37,823 to 41,006. Additional chromosome 19  
 map and sequence information are available at:  
<http://www-bio.liv.gov/bdrrp/genome/genome.html>.

FEATURES

source

1. 41006  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /map="19p13.1 between JAK3 and PDE4C"  
 /clone="R34383"  
 /cell\_line="5HL2-B"  
 /clone\_id="LL19NC03 R chromosome 19-specific cosmid  
 library"

COMMENT

note="Cosmid library constructed at LNL from flow-sorted  
 chromosomes from human-hamster hybrid 5HL2-B, which  
 carries chromosome 19 as its only human chromosome."  
 1. 67  
 /rpt\_family="AluJ/monomer"  
 143. .441  
 /rpt\_family="AluJb"  
 complement(578. 885)  
 /rpt\_family="AluSx"  
 complement(<1078. .1272)  
 note="LEYDIG INSULIN-LIKE PEPTIDE PRECURSOR"  
 /codon\_start=1

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

/product="INL3\_HUMAN [AA 1- 65]"  
/protein\_id="AAD22740.1"  
/db\_xref="GI:4558770"  
/translation="MDRLPAMALVLGLPALVFLGAPPTPEMKELCGHHFVALVR  
VCGGPMSTERRPATGDDRE"  
repeat\_region  
complement(1486, .1786)  
/rpt\_family="AluSx"  
repeat\_region  
complement(1803, .1970)  
/rpt\_family="AluSg/x"  
repeat\_region  
complement(1975, .2275)  
/rpt\_family="Aluub"  
repeat\_region  
complement(2279, .2444)  
/rpt\_family="Aluuo"  
repeat\_region  
complement(2447, .2747)  
/rpt\_family="AluSx"  
repeat\_region  
complement(2748, .2877)  
/rpt\_family="Aluuo"  
repeat\_region  
3052, .3357  
/rpt\_family="AluSx"  
repeat\_region  
complement(3365, .3412)  
/rpt\_family="MIR"  
repeat\_region  
complement(3773, .3846)  
/rpt\_family="(GA)n"  
repeat\_region  
complement(3851, .4142)  
/rpt\_family="AluSx"  
repeat\_region  
complement(4914, .4995)  
/rpt\_family="MIR"  
repeat\_region  
5386, .5686  
/rpt\_family="AluY"  
repeat\_region  
5739, .5848  
/rpt\_family="FLAM\_A"  
repeat\_region  
5870, .6170  
/rpt\_family="AluSx"  
repeat\_region  
complement(5918, .24239)  
/gene="JAK3"  
/note="JAK3 kinase 3 (a protein tyrosine kinase,  
leukocyte)"  
mRNA  
complement(join(5918, .6676, 9874, .9984, 10269, .10386,  
10994, .11166, 11440, .11564, 12285, .12474, 12556, .12695,  
14337, .14487, 14618, .14769, 14849, .14981, 15690, .15817,  
16895, .16979, 17698, .17829, 18029, .18156, 19243, .19429,  
19996, .20107, 21155, .21312, 21406, .21528, 22082, .22376,  
22793, .22938, 23146, .23257, 23543, .23666, 24000, .24239))  
/gene="JAK3"  
/product="Human JAK family protein tyrosine kinase (JAK3)  
mRNA, complete cds."  
repeat\_region  
complement(6416, .6472)  
/rpt\_family="MIR"  
repeat\_region  
complement(join(6509, .6676, 9874, .9984, 10269, .10386,  
10994, .11166, 11440, .11564, 12285, .12474, 12556, .12695,  
14337, .14487, 14618, .14769, 14849, .14981, 15690, .15817,  
16895, .16979, 17698, .17829, 18029, .18156, 19243, .19429,  
19996, .20107, 21155, .21312, 21406, .21528, 22082, .22376,  
22793, .22938, 23146, .23257, 23543, .23666, 24000, .24183))  
/gene="JAK3"  
/note="TYROSINE-PROTEIN KINASE JAK3; JANUS KINASE 3;  
JAK-3; LEUCOCYTE JANUS KINASE; L-JAK"  
/codon\_start=1  
/product="JAK3\_HUMAN"  
/protein\_id="AAD22741.1"  
/db\_xref="GI:4558771"  
/translation="MAPSEETPLIPQSCSLSTEGALHVLPPARGPPORLSFS  
FGDLADLCVOAKKASGITLVHSLPALATEDLSCEPSPSHFSVEDASTOYLVI  
RFEPNMFLEKCHRGFLRKDLASAIIDLPLEHLRFQHSRDLSCGLPYGLSLKEGG  
ECSLAVLDLARBARBOAORPGELTKVSKACLPRLRDLGSLSVTRRRIRRYR  
RALRVAAACDARSLMAKYIMDERIDPAGAAFTFVIGPGALGSGDGLLVAAD  
GGIATVGEBOEVLOPCDFPEIVDITSKQAPRPGAEHRIVYTRDNOILKEPFG  
LPEALSFVALVDGEFRLTJTDSDHCEKVEAPARGLLEVAQCHPITLDAIINKLTKG  
GSRGSYVLRSSPDSPFLTYCVQNPGLDYGKGLIRSPGTFLLVGSRRHSL  
RELATCMGDLGHVDAVATLTSCCIPRREKSNLIYVGRSHSPSTSLVOPOSYOY  
SOMTFKIPADSLFEMHENGHSFTKTYRGCRHEVVDGAPKTFVILLKVDARKKNM  
ESPLEASLMSQVSYRHLVLLHGYCMAGDSTMOEFPHLGAIIDMYLEKRRHLPASK  
LQYVKQLATVLTLEDKGLPHGNVSAKKVLLARLGAGDGSPEFTKLSDPGVSPALVSL

repeat\_region  
complement(6747, .7040)  
/rpt\_family="AluSx"  
repeat\_region  
complement(7111, .7412)  
/rpt\_family="AluSg"  
repeat\_region  
7441, .7688  
/rpt\_family="AluSp"  
repeat\_region  
complement(7985, .8020)  
/rpt\_family="(CAAA)n"  
repeat\_region  
complement(8021, .8315)  
/rpt\_family="Aluub"  
repeat\_region  
complement(8316, .8454)  
/rpt\_family="Aluub"  
repeat\_region  
8499, .8564  
/rpt\_family="LINE2"  
repeat\_region  
8643, .8902  
/rpt\_family="Aluuo"  
repeat\_region  
complement(8929, .9224)  
/rpt\_family="AluSx"  
repeat\_region  
9243, .9536  
/rpt\_family="Aluuo"  
repeat\_region  
complement(9560, .9753)  
/rpt\_family="MIR"  
repeat\_region  
11759, .12044  
/rpt\_family="AluSx"  
repeat\_region  
complement(12074, .12154)  
/rpt\_family="MIR"  
repeat\_region  
13194, .13323  
/rpt\_family="FLAM\_A"  
repeat\_region  
13337, .13638  
/rpt\_family="AluSx"  
repeat\_region  
complement(13651, .13805)  
/rpt\_family="Aluub"  
repeat\_region  
13832, .14140  
/rpt\_family="AluSx"  
repeat\_region  
complement(16047, .16337)  
/rpt\_family="AluSx"  
repeat\_region  
complement(17180, .17480)  
/rpt\_family="AluSx"  
repeat\_region  
complement(17487, .17609)  
/rpt\_family="FLAM\_C"  
repeat\_region  
complement(17910, .17976)  
/rpt\_family="MIR"  
repeat\_region  
complement(18301, .18479)  
/rpt\_family="AluSg/x"  
repeat\_region  
complement(18480, .18780)  
/rpt\_family="AluSx"  
repeat\_region  
complement(19505, .19804)  
/rpt\_family="AluSc"  
repeat\_region  
20248, .20548  
/rpt\_family="AluY"  
repeat\_region  
complement(20590, .20882)  
/rpt\_family="Aluuo"  
repeat\_region  
23362, .23462  
/rpt\_family="MIR"  
repeat\_region  
complement(24063, .24094)  
/rpt\_family="GC-rich"  
repeat\_region  
24283, .24582  
/rpt\_family="AluSx"  
repeat\_region  
24583, .24608  
/rpt\_family="(GAA)n"  
repeat\_region  
24851, .24930  
/rpt\_family="AluJ/FLAM"  
repeat\_region  
complement(24932, .25220)  
/rpt\_family="AluSx"  
repeat\_region  
25221, .25284

MLTDRIEMVAPECELEAOTLSLEADKMGFGATVWEVSGTMTPIASLDPAKKLOFYED  
RQOLPARKWTELALLIOCMAYEPVOPSPFPAVIRDLNLSISSPYELLSDPTPALAP  
RDGLMNAQLYACDDPTIFERRHKITISQLEKMGFGVELCRYPDLDDNGALVAVQ  
LOHSGPDQGRFOREIOITLKAHSDFTVKTGVSFGKQSLRLVMEYLPSCGLRDL  
QRRHARLADASRLLLYSSQICKMEYLSRCVHRDLAARNILVSEAHVIAIDGLAK  
LLPLDKDYVYVREPGOSPFIWAPESLSDNIFSRQSPVWSGVVLYELFYCDKSCSP  
SAEFLRMWGCERDVPAALCRLELEEGORLPAAPACPAEVEHLMKLCMAVSPQDRPSF  
SALGPOLDMLWSGSRGCEHTAFTHPGKHHSLSFS"  
repeat\_region  
complement(6747, .7040)  
/rpt\_family="AluSx"  
repeat\_region  
complement(7111, .7412)  
/rpt\_family="AluSg"  
repeat\_region  
7441, .7688  
/rpt\_family="AluSp"  
repeat\_region  
complement(7985, .8020)  
/rpt\_family="(CAAA)n"  
repeat\_region  
complement(8021, .8315)  
/rpt\_family="Aluub"  
repeat\_region  
complement(8316, .8454)  
/rpt\_family="Aluub"  
repeat\_region  
8499, .8564  
/rpt\_family="LINE2"  
repeat\_region  
8643, .8902  
/rpt\_family="Aluuo"  
repeat\_region  
complement(8929, .9224)  
/rpt\_family="AluSx"  
repeat\_region  
9243, .9536  
/rpt\_family="Aluuo"  
repeat\_region  
complement(9560, .9753)  
/rpt\_family="MIR"  
repeat\_region  
11759, .12044  
/rpt\_family="AluSx"  
repeat\_region  
complement(12074, .12154)  
/rpt\_family="MIR"  
repeat\_region  
13194, .13323  
/rpt\_family="FLAM\_A"  
repeat\_region  
13337, .13638  
/rpt\_family="AluSx"  
repeat\_region  
complement(13651, .13805)  
/rpt\_family="Aluub"  
repeat\_region  
13832, .14140  
/rpt\_family="AluSx"  
repeat\_region  
complement(16047, .16337)  
/rpt\_family="AluSx"  
repeat\_region  
complement(17180, .17480)  
/rpt\_family="AluSx"  
repeat\_region  
complement(17487, .17609)  
/rpt\_family="FLAM\_C"  
repeat\_region  
complement(17910, .17976)  
/rpt\_family="MIR"  
repeat\_region  
complement(18301, .18479)  
/rpt\_family="AluSg/x"  
repeat\_region  
complement(18480, .18780)  
/rpt\_family="AluSx"  
repeat\_region  
complement(19505, .19804)  
/rpt\_family="AluSc"  
repeat\_region  
20248, .20548  
/rpt\_family="AluY"  
repeat\_region  
complement(20590, .20882)  
/rpt\_family="Aluuo"  
repeat\_region  
23362, .23462  
/rpt\_family="MIR"  
repeat\_region  
complement(24063, .24094)  
/rpt\_family="GC-rich"  
repeat\_region  
24283, .24582  
/rpt\_family="AluSx"  
repeat\_region  
24583, .24608  
/rpt\_family="(GAA)n"  
repeat\_region  
24851, .24930  
/rpt\_family="AluJ/FLAM"  
repeat\_region  
complement(24932, .25220)  
/rpt\_family="AluSx"  
repeat\_region  
25221, .25284

```

/rpt_family="AluJo/FLAM"

Alignment Scores:
Pred. No.: 1.33e-26 Length: 41006
Score: 39.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.55% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AC007201 (1-41006)

QY 989 TyralProGluSerLeuSerAspAsnIlePheSerArgInSerAspValTrpSerPhe 1008
Db 10385 TATGCCCGCATGCTCTCGACACATCTCTCTCGCCGCTCAGACGCTTGAGAGCTTC 10326

QY 1009 GlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027
Db 10325 GGGGTGCTCTGTACGAGCTCTTCACCTACTGCGACAAAGCTCGAGCCCTCGGCC 10269

RESULT 19
LOCUS HSU08340 198 bp mRNA linear PRI 21-APR-1994
DEFINITION Human clone NTK16 tyrosine kinase mRNA, partial cds.
ACCESSION U08340
VERSION U08340.1 GI:473879
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 198)
AUTHORS Fuortes M.
TITLE Tumor Necrosis Factor and Adhesion: Joint Control of Neutrophil
JOURNAL Activation
Tnesis (1994) Cell Biology and Anatomy, Cornell University,
Graduate School of Medical Sciences
2 (bases 1 to 198)
AUTHORS Fuortes M.
TITLE Direct Submission
JOURNAL Submitted (05-APR-1994) Michele Fuortes, Cell Biology and
Anatomy/Medicine, Cornell University Medical College, 1300 York
Avenue, New York, NY 10021, USA

FEATURES
source
1. 198
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NTK16"
/cell_type="neutrophil"
/tissue_type="blood"
<1..>198
/codon_start=1
/evidence-not-experimental
/product="tyrosine kinase"
/protein_id="AA17743.1"
/db_xref="GI:473880"
/translation="DLARNIVLESEAHVKIADFGIAKLPLDLDKYVVEPQSPILF
WYAPESIDNIFSRSDVMSFG"

BASE COUNT 41 a 66 c 52 g 39 t
ORIGIN

Alignment Scores:
Pred. No.: 1.66e-26 Length: 198
Score: 37.00 Matches: 65
Percent Similarity: 97.01% Conservative: 0
Best Local Similarity: 97.01% Mismatches: 1
Query Match: 3.37% Indels: 2
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x HSU08340 (1-198)

QY 944 AspleuAlaAaGAsnIleLeuValGluSerGluAlaHisValLysIleAlaAspPhe 963
|||||

```

```

Db 1 GACCTGGCGCCCGCAACATCTCTGTGAGAGCGAGCCACACGTCAAGATCGTCACTTC 60

QY 964 GlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrValValArgGluProGly 983
Db 61 GGCTTACTAAGCTGCTGCGCCCTTGAC-AAAGACTACTACTGCTGCGAGCCAGGCCA 119

QY 983 nserProIlePheThrPyrAlaProGluSerLeuSerAspAsnIlePheSerArgInse 1003
Db 120 GAGCCCATTTTCTGTGATGCGCCCGAATCCCTCTCGGACACATCTTCTCGCCAGTC 179

QY 1003 rAspValTrpSerPheGly 1009
Db 180 AGACGTGTGAGCTTCGGA 198

RESULT 20
LOCUS E08798 165 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding novel human tyrosine kinase.
ACCESSION E08798
VERSION E08798.1 GI:2176910
KEYWORDS JP 1995059569-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 165)
AUTHORS Sakano S.
TITLE NOVEL TYROSINE KINASE PEPTIDE AND DNA CODING THE SAME
JOURNAL Patent: JP 1995059569-A 1 07-MAR-1995;
ASAHI CHEM IND CO LTD
OS Homo sapiens (human)
PN JP 1995059569-A/1
PD 07-MAR-1995
PF 25-AUG-1993-EP-1993210404
PI SAKANO SEIJI
PC C12N15/09,C12N9/12;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FT source 1. 165
/organism="Homo sapiens"
/cell_type="megakaryoblast"
/cell_line="U9-7"
FT mat_peptide 1. 165
/product="novel human tyrosine kinase"
FT Location/Qualifiers
1. 165
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 36 a 56 c 39 g 34 t
ORIGIN

Alignment Scores:
Pred. No.: 1.01e-20 Length: 165
Score: 31.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.82% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x E08798 (1-165)

QY 973 LysAspTyrTyrValValArgGluProGlyGlnSerProIlePheThrPyrAlaProGlu 992
Db 73 AAAGACTACTAAGCTGCTGCGCCCGAATCCCTCTCGGACACATCTTCTGTGATGCGCCGAA 132

QY 993 SerLeuSerAspAsnIlePheSerArgInser 1003
|||||
Db 133 TCCTCTGCGACACATCTTCTCTCGCCAGTCA 165
|||||

```

```

RESULT 21
LOCUS AF034576 3644 bp mRNA linear VRT 03-SEP-1998
DEFINITION Gallus gallus Janus tyrosine kinase (JAK) mRNA, complete cds.
ACCESSION AF034576
VERSION AF034576.1 GI:2645986
KEYWORDS
SOURCE
ORGANISM Gallus gallus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 3644)
AUTHORS Sofer, L., Kampa, D., and Burnside, J.
TITLE Molecular cloning of a chicken JAK homolog from activated T cells.
JOURNAL Gene 215 (1), 29-36 (1998)
MEDLINE 98332716
PUBMED 9666067
REFERENCE 2 (bases 1 to 3644)
AUTHORS Sofer, L., Kampa, D., and Burnside, J.
TITLE Direct Submission
JOURNAL Submitted (14-NOV-1997) Animal Science, University of Delaware, 40 Townsend Hall, Newark, DE 19716, USA
FEATURES
source
1..3644
location/Qualifiers
/organism="Gallus gallus"
/sub_species="domesticus"
/db_xref="taxon:9031"
/cell_type="lymphocyte"
/tissue_type="spleen"
/notes="cells activated with concanavalin A"
1..3644
/gene="JAK"
91..3411
/gene="JAK"
/feature="JAK"
/feature="similar to JAK protein kinase"
/codon_start=1
/product="Janus tyrosine kinase"
/protein_id="AAC34195.1"
/db_xref="GI:2645986"
translation="MAPLGEETPLIGRSCSISSEPTGLVLYTHRGPHAPPSAAT
LTFEGEYAEELCYHAKACGVLPICHPALATEDISCMYPRHPTFTVADASOY
VYRIFFEPNMGOGVHROPSPDRPSPVADIPYDILFPOSSDPLFAGMELALSL
AGOEELSLAVLDMLRIKERMOSKPEFVSQVKTCTIPEDLRQIOHSLTKRIR
RVAOSLRMGSCRYDCCCLAKTILDLERLCCAESFHAHGDADIAIHVETDSG
VSMSCVGSERQHCDFPDIDAVSITKQSRGQVEMNIVLTITDNVLEVEPTLR
EALSFMAVLDGYRLTTADAHYFCKEVAPRLLEDNMOCHGISSEFAVRKLAAGS
HPGLVLRSPQDESDYLLTYCAETRSQDYKRLIRDEGGSFMLGIARFCSLOE
LGTGCGGLAEGAHRLDTCPCPLPREKSNLLIVRSGCPRPNSPPAPRSPNOSL
FKHIDPESLIGESLGGSTFTHYKIKIRKDDPEVOTPVYLVKVMSSHRNCSSEL
EASISMSLSHKLHYLLHGVSLGKDSIMVQETIRHPLDLYLKNHSEGVTTSMKIQ
VAKOLATALNLEDKKTTGNVSAKKVLTREGDAASSPFIKLNPGVSTTVLAKE
WLVERIPVNAPECLSDQSLALPADKQFGATLWEITSGMPPSLPEPKQLQFYS
RIOLPAPMSLALIAQCIAPSRPFRALINDISLSDYELGSLSELSPGVVTR
ESGCGYEHVAGHGPAPPEERHLYKISLKGNGSVELICRYDPLDSTGTLAVKTL
QODSAKLEHRTLLYAMOCKMEYLGARQCHRDALNSVETSHVKIGDFELANL
LPQDKDYVVOPEQSPVFWYAPASRLADNVSFASDVSFVLLYELFTYSNKRSPS
EFLHMEPEKPAQIICHLLELLLDLSRLPVPCCPMIEVYAMLSQWAFASAPRTPT
ELARVLELRGRGRTANG"
BASE COUNT 697 a 1173 c 1150 g 624 t
ORIGIN

```

## Alignment Scores:

```

Pred. No.: 1.14e-18 Length: 3644
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.73% Indels: 0
DB: 5 Gaps: 0

```

US-09-397-967-16 (1-1099) x AF034576 (1-3644)

```

Oy 813 PhcIUGlUArghHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPhcIyseryl 832
Db 2536 TTCGAGAGAGAGGACCTCAAGTACATCTGCTGGCAAGGCAACTTGGAGCGTG 2595
Oy 833 GluLeuCyArgTyrAspProLeuGlyAsp 842
Db 2596 GACCTGTGCCCTACGACCCGCTGGGTGAC 2625
RESULT 22
LOCUS AX203088 811 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 9 from Patent WO0152892.
ACCESSION AX203088
VERSION AX203088.1 GI:15392446
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS Vastis, G.
TITLE Jak/stat pathway inhibitors and the uses thereof
JOURNAL Patent: WO 0152892-A 9 26-JUL-2001;
GENZYME CORPORATION (US)
FEATURES
source
1..811
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 149 a 254 c 226 g 182 t
ORIGIN

```

## Alignment Scores:

```

Pred. No.: 3.22e-18 Length: 811
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 6 Gaps: 0

```

US-09-397-967-16 (1-1099) x AX203088 (1-811)

```

Oy 1006 TRPSPRPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerPro 1025
Db 168 TGGACCTTCGGGGGCGTCTGTACAGACTCTTACCTGACGACAAAGCTGACGCC 227
Oy 1026 SerIaGluPheLeuArgMetMetGly 1034
Db 228 TCGGCGGAGTTCTGCGGATGATGGGA 254

```

```

RESULT 23
LOCUS AC103459 159623 bp DNA linear HTG 12-JUL-2002
DEFINITION Rattus norvegicus clone CH230-30H1, *** SEQUENCING IN PROGRESS ***
ACCESSION AC103459
VERSION AC103459.3 GI:21728723
KEYWORDS HTG; HTGS; PHASEL.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 159623)

```

```

REFERENCE
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., All-rosman, F.R., Allen, C.,
Alsbrooks, S.L., Amaralunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barberia, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowles, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, R.L., Byrd, N.C.,
Carion, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,

```



```

* 58819 58918: gap of unknown length
* 58919 60329: contig of 141 bp in length
* 60330 60429: gap of unknown length
* 60430 61637: contig of 1208 bp in length
* 61638 61737: gap of unknown length
* 61738 63760: contig of 2023 bp in length
* 63761 63860: gap of unknown length
* 63861 65870: contig of 2010 bp in length
* 65871 65970: gap of unknown length
* 65971 67818: contig of 1848 bp in length
* 67819 67918: gap of unknown length
* 67919 70369: contig of 2451 bp in length
* 70370 70469: gap of unknown length
* 70470 73659: contig of 3190 bp in length
* 73660 73759: gap of unknown length
* 73760 76872: contig of 3113 bp in length
* 76873 76972: gap of unknown length
* 76973 79979: contig of 3007 bp in length
* 79980 80079: gap of unknown length
* 80080 83397: contig of 3318 bp in length
* 83398 83497: gap of unknown length
* 83498 86200: contig of 2703 bp in length
* 86201 86300: gap of unknown length
* 86301 89469: contig of 3169 bp in length
* 89470 89569: gap of unknown length
* 89570 91678: contig of 2109 bp in length
* 91679 91778: gap of unknown length
* 91779 95960: contig of 4182 bp in length
* 95961 96060: gap of unknown length
* 96061 98742: contig of 2682 bp in length
* 98743 98842: gap of unknown length
* 98843 101456: contig of 2614 bp in length

```

## Alignment Scores:

```

Pred. No.: 2,2e-16 Length: 159623
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

```

US-09-397-967-16 (1-1099) x AC103459 (1-159623)

QY 563 SerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeu 582  
 |||||||

DB 52329 TCCTTCTGAGAGCCGACGCTTGTGACCAAGTATCTACCCGACCTGGTGTCTG 52388  
 |||||||

QY 583 HisGlyValCysMetAlaGlyAspSer 591  
 |||||||

DB 52389 CATGAGTCTGCATGCTGAGACACT 52415  
 |||||||

## RESULT 24

BC028068 2881 bp mRNA linear PRI 01-MAY-2002

DEFINITION Homo sapiens, similar to Janus kinase 3 (a protein tyrosine kinase, leukocyte), clone MGC:39993 IMAGE:5212575, mRNA, complete cds.

## ACCESSION

BC028068

## VERSION

BC028068.1

## KEYWORDS

MGC.

## SOURCE

Homo sapiens.

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

Strausberg, R.

Direct Submission

Submitted (08-APR-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nrl.nih.gov](mailto:nisc_mgc@nrl.nih.gov)  
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Latic, P., Legaspi, R., Maduro, Q.L.,  
 Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J.,  
 Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C.,  
 Voigt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A.,  
 Zhang, L.-H. and Green, E.D.

## FEATURES

## CDS

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 62 Row: 9 Column: 8  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 4557680.  
 Location/Qualifiers  
 1..2881  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:39993 IMAGE:5212575"  
 /tissue\_type="blood, adult leukocytes"  
 /clone\_id="NIH\_MGC\_118"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SF0RT6"  
 91..1950  
 /codon\_start=1  
 /product="Similar to Janus kinase 3 (a protein tyrosine  
 kinase, leukocyte)"  
 /protein\_id="AAH28068.1"  
 /db\_xref="gi:20380125"  
 /translation="MAPPSEETPLIPRSCSLSTEGALHVLHPARGPGPQRLSFS  
 FGDHLAEDLCVOAKAGSGLIPYHSLFPLALEDSCPPSPHFEVDASTQVLYRI  
 RFLFPMFGLKCHRFGRGLDLASAILDLPVLEHFAQHSOLVSGRLPVGLSLTEOG  
 ECLSLAVDLIARMAREOARQBELIKTVSNACLPSPSLRDLQGLSPFTRRRIRTRV  
 RALRRVACQADRHSLMAKYINDLERLPAGAAEFPHYGLPGLAGHGLGILRPACD  
 GGIAVTSGEDEVLPFCPEIVDLSIQAPRVGAGHRLVTVTRTNOQILEAFPG  
 LPEALSFVALVDGYFRLLTDSQHFCKEVAAPRLLEVAEOCHGPIITDFAINKLKTG  
 GSRPGSVYLRSPDPDSFLTLVCQVNPGLPKYKCLLRSPSTFTFLVLSGRSPSYL  
 RELATVCMGDGLHVDGVAVLTSCCIPRKEKSNLIVORSGHPTSSIVQPOSOYOL  
 SOMPEFKIPADSLKEHNIHGHSFTKTYRGCHHEVVDGAEARTEVLLKVMKKHNCM  
 ESFLFAALMSQVSYRHLVLLHGVCMADDSPPPTHTPTPASPKSLRPLPLF"

## BASE COUNT

595 a 867 c 797 g 602 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 8,37e-17 Length: 2881  
 Score: 28.00 Matches: 28  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.55% Indels: 0  
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x BC028068 (1-2881)

QY 173 MetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAla 192  
 |||||||

DB 607 ATGCGCGAG 666  
 |||||||

QY 193 CysLeuProProSerLeuArgasp 200  
 |||||||

DB 667 TGCTTACCCCAAGCCTGCGGAC 690  
 |||||||

## RESULT 25

AX203082

LOCUS AX203082 109 bp DNA linear PAT 30-AUG-2001

DEFINITION Sequence 3 from Patent WO0152892.

ACCESSION AX203082  
VERSION AX203082.1 GI:15392439  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
TITLE Jak/stat pathway inhibitors and the uses thereof  
JOURNAL Patent: WO 0152892-A 3 26-JUL-2001;  
GENZYME CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1.109  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 22 a 36 c 34 g 15 t 2 others  
ORIGIN

Alignment Scores:  
Pred. No.: 3.4e-08 Length: 109  
Score: 18.00 Matches: 18  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.64% Indels: 0  
DB: Gaps: 0

US-09-397-967-16 (1-1099) x AX203082 (1-109)

QY 945 LeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValIleAlaasp 962  
|||||  
56 TTGGCCGCCGAAACATCTCTGGAGAGCAGCAGCAGCAGATGCTGAC 109

RESULT 26  
LOCUS AX203085 726 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 6 from Patent W00152892.  
ACCESSION AX203085  
VERSION AX203085.1 GI:15392442  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
TITLE Jak/stat pathway inhibitors and the uses thereof  
JOURNAL Patent: WO 0152892-A 6 26-JUL-2001;  
GENZYME CORPORATION (US)  
FEATURES  
source Location/Qualifiers  
1.726  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 151 a 224 c 204 g 147 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.55e-07 Length: 726  
Score: 18.00 Matches: 18  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.64% Indels: 0  
DB: Gaps: 0

US-09-397-967-16 (1-1099) x AX203085 (1-726)

QY 559 AsnGlyMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyr 576  
|||||  
Db 635 AACGCAATGAGATCTCTGGAAGCAGCAGCTGATGACCAAGTCTCTAC 688

RESULT 27  
AF227557  
LOCUS AF227557 495 bp mRNA linear PRI 01-JUN-2000

DEFINITION Macaca mulatta Janus kinase-3 (JAK3) mRNA, partial sequence.  
AF227557  
VERSION AF227557.1 GI:8132806  
KEYWORDS  
SOURCE Macaca mulatta.  
ORGANISM Macaca mulatta  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.  
TITLE Cytokine Signal Transduction Genes from Rhesus Macaques  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 495)  
REFERENCE Arredondo, J.  
TITLE Direct Submission  
JOURNAL Submitted (24-JAN-2000) California Regional Primate Research  
Center, University of California-Davis, One Shields Avenue, Davis,  
CA 95616, USA  
FEATURES  
source Location/Qualifiers  
1.495  
/organism="Macaca mulatta"  
/db\_xref="taxon:9544"  
/cell\_type="peripheral blood mononuclear cells"  
gene <1..>495  
/gene="JAK3"  
misc\_feature <1..>495  
/note="similar to Janus kinase-3; tyrosine kinase class I  
receptor; contains frameshifts"  
BASE COUNT 100 a 157 c 132 g 106 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.08e-06 Length: 495  
Score: 17.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.55% Indels: 0  
DB: Gaps: 0

US-09-397-967-16 (1-1099) x AF227557 (1-495)

QY 514 AspSerLeuGluTrpHisGluAsnLeuGlyHisGlySerPheTrpIle 530  
|||||  
Db 376 GACAGCCTAGAGTGCATGAGACACTGGCCATGATCTTCACCAAGATT 426

RESULT 28  
LOCUS HSU57096 466 bp mRNA linear PRI 18-JUN-1996  
DEFINITION Human Janus kinase 3 (JAK3) mRNA, partial cds.  
ACCESSION U57096  
VERSION U57096.1 GI:1354824  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
TITLE Expression of Janus kinase 3 in human endothelial and other  
non-lymphoid and non-myeloid cells  
JOURNAL Verbsky, J.W., Bach, E.A., Fang, Y.F., Yang, L., Randall, D.A. and  
J. Biol. Chem. 271 (24), 13976-13980 (1996)  
MEDLINE 96278845  
PUBMED 8662778  
REFERENCE 2 (bases 1 to 466)  
AUTHORS Fields, L.E.  
TITLE Direct Submission  
JOURNAL Submitted (30-APR-1996) Larry E. Fields, Medicine, Washington  
University School of Medicine, 660 S. Euclid Ave., Campus Box 8086,  
St. Louis, MO 63110-1093, USA  
FEATURES  
source Location/Qualifiers



source 1..466  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="vascular endothelial cells (HUVEC) obtained from Clonetics Corporation, San Diego, CA"  
gene 1..466  
/gene="jak3"  
<1..>466  
/db\_xref="GI:1354825"  
CDS /feature="phosphorylates Stat6, IRS1, IRS2, PI3K"  
/note="protein tyrosine kinase"  
/codon\_start=1  
/product="janus kinase 3"  
/protein\_id="AAC50542.1"  
/db\_xref="GI:1354825"  
/translation="GPGPRLSFSFGDHLAEDLCVQAAKASGILPVYHSPLATEDLS  
CWPPSHIFSVEDASTOVLLIRIRFPNPFGLKCRFGRLKDLASLIDLPVLEHL  
FAQHRSDLVGRLPYGLSLKEGELSLAVLDLARMARQAOQRPGLLKAVSYK"  
BASE COUNT 70 a 153 c 143 g 100 t  
ORIGIN

Alignment Scores:  
Pred. No.: 9.68e-06 Length: 466  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x HS057096 (1-466)

Oy 62 Gly1LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAsp 77  
Db 79 GGCAATCGCTGGCTGTACCACTCCTTGTCTGCGCAGCAGAGAC 126

RESULT 29  
DRE5691 1967 bp mRNA linear VRT 30-SEP-1998  
LOCUS DRE5691 1967 bp mRNA linear VRT 30-SEP-1998  
DEFINITION Danilo rerio protein tyrosine kinase (jak2b), partial.  
ACCESSION AJ005691  
VERSION AJ005691.1 GI:3687399  
KEYWORDS jak2b gene; protein tyrosine kinase; signal transduction.  
SOURCE zebrafish.  
ORGANISM Danilo rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danilo.  
1 (bases 1 to 1967)  
Oates, A.C.  
Direct Submission  
Submitted (22-APR-1998) Oates A.C., Growth Regulation, Ludwig  
Institute for Cancer Research, PO Royal Melbourne Hospital,  
Parkville, VIC, 3050, AUSTRALIA  
2 (bases 1 to 1967)  
Oates, A.C.  
Thesis (1998) University of Melbourne, Australia  
Location/Qualifiers  
1..1967  
/organism="Danilo rerio"  
/db\_xref="taxon:7955"  
1..1497  
/gene="jak2b"  
<1..1497  
/gene="jak2b"  
/EC\_number="2.7.1.12"  
/function="signal transduction"  
/codon\_start=1  
/product="protein tyrosine kinase"  
/protein\_id="CAA0675.1"  
/db\_xref="GI:3687400"  
/db\_xref="SPTREMBL:O93597"  
/translation="SVNIIWKLEVAQKOLANAMLYLEEKSLAHGVCAKNIILLIREEDR  
ALGNTPIKISDPCISITVLPRLIVERIRFWVPPECILDPKKNLSLADKMSFGTTLME

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES

ICSGEOPLIAMNDSSKKHLPEYNNHOLPAKWTFLANLINSCHMEPEPTPRSKAIR  
DLNLFCEPDYIVKESDIMPSSRAASTFNTGTFRNNEVQEEERHLIFLOQLGKNG  
SVEMCRDPLDNDNGEYVAVKRLKQSTTEHTRDEREIEILKSLOHNIYVKGCYQ  
AGRNLRLVMEYELPYGSLRDYLNKRNPRIDHOKLVHASQCKMEYLAIRYIHRDL  
ATRNILVESECRVKIGDEGLTKVLPQKREKVPKPEGSPLEFWAPSPSTSKSVAS  
DVSEGVLYELFTYSDKICSPPTVLSMGKDGOGOTIVYHLIELKRGRLRPPQC  
CPTMEFEMDCMNDPSLRNREKELARDLINDSDADRYTQVPER"  
BASE COUNT 563 a 429 c 498 g 476 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 3.06e-05 Length: 1967  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x DRE5691 (1-1967)

Oy 1003 SerAspValTyrSerPheGlyValLeuTyrGluLeuPheThrTyr 1018  
Db 1174 TCCGACGTGTGGAGTTTGGAGTGGTCTGTATGAACCTTTACATAC 1221

RESULT 30  
ARI29823 3435 bp DNA linear PAT 16-MAY-2001  
LOCUS ARI29823 3435 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 21 from patent US 6187552.  
ACCESSION ARI29823  
VERSION ARI29823.1 GI:14117720  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3435)  
AUTHORS Robert, S.L. and Kaytes, P.S.  
TITLE Method for identifying inhibitors of JAK2/cytokine receptor binding  
JOURNAL Patent: US 6187552-A 21 13-FEB-2001;  
FEATURES Location/Qualifiers  
1..3435  
/organism="unknown"  
BASE COUNT 1157 a 598 c 733 g 947 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.78e-05 Length: 3435  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x ARI29823 (1-3435)

Oy 1003 SerAspValTyrSerPheGlyValLeuTyrGluLeuPheThrTyr 1018  
Db 3106 TCAGATGTTGGAGCTTGGAGTGGTCTGTATGAACCTTTACATAC 3153

RESULT 31  
AR048200 3495 bp DNA linear PAT 29-SEP-1999  
LOCUS AR048200 3495 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 2 from patent US 5821069.  
ACCESSION AR048200  
VERSION AR048200.1 GI:5970543  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3495)  
AUTHORS Wilks, A. Frederick, J. Ziemlecki, A. and Harpur, A.  
TITLE Method for determining tyrosine kinase in a sample  
JOURNAL Patent: US 5821069-A 2 13-OCT-1998;  
FEATURES Location/Qualifiers

```

source 1..3495
/organism="unknown"
BASE COUNT 1121 a 688 c 794 g 892 t
ORIGIN

Alignment Scores:
Pred. No.: 4.85e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AR048200 (1-3495)

OY 1003 SeraspValITPserPheGlyValValLeuTYrGluLeuPheThTYr 1018
|||||
Db 2695 TCAGATGTGTGAGCTTGGAGTGTCTCTATACGAACCTTTCACATAC 2742

RESULT 32
LOCUS AR068116 3495 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5852184.
ACCESSION AR068116
VERSION AR068116.1 GI:5999338
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3495)
AUTHORS Wilks/A.Frederick. and Harpur/A.
TITLE Protein tyrosine kinase
JOURNAL Patent: US 5852184-A 2 22-DEC-1998;
FEATURES
Source 1..3495
Location/Qualifiers
/organism="unknown"

BASE COUNT 1121 a 688 c 794 g 892 t
ORIGIN

Alignment Scores:
Pred. No.: 4.85e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AR068116 (1-3495)

OY 1003 SeraspValITPserPheGlyValValLeuTYrGluLeuPheThTYr 1018
|||||
Db 2695 TCAGATGTGTGAGCTTGGAGTGTCTCTATACGAACCTTTCACATAC 2742

RESULT 33
LOCUS AR071176 3495 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 2 from patent US 5910426.
ACCESSION AR071176
VERSION AR071176.1 GI:7222064
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3495)
AUTHORS Wilks/A.Frederick., Ziemiecki/A. and Harpur/A.
TITLE Protein tyrosine kinase
JOURNAL Patent: US 5910426-A 2 08-JUN-1999;
FEATURES
Source 1..3495
Location/Qualifiers
/organism="unknown"

BASE COUNT 1121 a 688 c 794 g 892 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 4.85e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AR071176 (1-3495)

OY 1003 SeraspValITPserPheGlyValValLeuTYrGluLeuPheThTYr 1018
|||||
Db 2695 TCAGATGTGTGAGCTTGGAGTGTCTCTATACGAACCTTTCACATAC 2742

RESULT 34
LOCUS AR177639 3495 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 3 from patent US 6312941.
ACCESSION AR177639
VERSION AR177639.1 GI:17919994
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3495)
AUTHORS Carter-Su,C., Rul,L.,Y. and Karow,D.S.
TITLE Compositions and methods for identifying signaling pathway agonists
and antagonists
JOURNAL Patent: US 6312941-A 3 06-NOV-2001;
FEATURES
Source 1..3495
Location/Qualifiers
/organism="unknown"

BASE COUNT 1121 a 688 c 794 g 892 t
ORIGIN

Alignment Scores:
Pred. No.: 4.85e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AR177639 (1-3495)

OY 1003 SeraspValITPserPheGlyValValLeuTYrGluLeuPheThTYr 1018
|||||
Db 2695 TCAGATGTGTGAGCTTGGAGTGTCTCTATACGAACCTTTCACATAC 2742

RESULT 35
LOCUS I62349 3495 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 2 from patent US 5658791.
ACCESSION I62349
VERSION I62349.1 GI:2480297
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3495)
AUTHORS Wilks/A.Frederick., Ziemiecki/A. and Harpur/A.
TITLE Antibodies which specifically bind to proteins having tyrosine
kinase activity, wherein said proteins have more than one tyrosine
kinase domain, and no SH2 domains
JOURNAL Patent: US 5658791-A 2 19-AUG-1997;
FEATURES
Source 1..3495
Location/Qualifiers
/organism="unknown"

BASE COUNT 1121 a 688 c 794 g 892 t
ORIGIN

Alignment Scores:
Pred. No.: 4.85e-05 Length: 3495

```

Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x 162349 (1-3495)

QY 1003 SerAspValTrpSerPheGlyValValLeuTYRGluLeuPheThrTYR 1018  
DB 2695 TCAGATGTGTGGAGCTTTGGAGCTTCTATACGAACCTTTCACATAC 2742

RESULT 36

LOCUS 187981 3495 bp DNA linear PAT 10-AUG-1998  
DEFINITION Sequence 2 from patent US 5716818.  
ACCESSION 187981  
VERSION 187981.1 GI:3407921  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 3495)  
AUTHORS Wilks,A.Frederick., Ziemlecki,A. and Harpur,A.  
TITLE Protein tyrosine kinase  
JOURNAL Patent: US 5716818-A 2 10-FEB-1998;  
FEATURES Location/Qualifiers  
source 1..3495  
/organism="unknown"

BASE COUNT 1121 a 688 c 794 g 892 t  
ORIGIN

Alignment Scores:

Pred. No.: 4.85e-05 Length: 3495  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x 187981 (1-3495)

QY 1003 SerAspValTrpSerPheGlyValValLeuTYRGluLeuPheThrTYR 1018  
DB 2695 TCAGATGTGTGGAGCTTTGGAGCTTCTATACGAACCTTTCACATAC 2742

RESULT 37

LOCUS AF001362 3500 bp mRNA linear PRI 17-JUN-1998  
DEFINITION Homo sapiens Jak2 kinase (JAK2) mRNA, complete cds.  
ACCESSION AF001362  
VERSION AF001362.1 GI:3236321  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 3500)  
AUTHORS Datal,I., Arpaia,E., Dadi,H., Kulkarinl,S., Squire,J. and  
Rolfman,C.M.  
TITLE Cloning and characterization of the human homolog of mouse Jak2  
JOURNAL Blood 91 (3), 844-851 (1998)  
MEDLINE 98118198  
PUBMED 9446644  
REFERENCE 2 (bases 1 to 3500)  
AUTHORS Rolfman,C.M.  
TITLE Direct Submission  
JOURNAL Submitted (26-APR-1997) Immunology, Hospital for Sick Children, 555  
University Ave., Toronto, Ontario M5G 1X8, Canada  
FEATURES Location/Qualifiers  
source 1..3500  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

/Chromosome="9"  
/map="p23-24"  
/cell\_line="pre-B leukaemia cell line G2"  
/cell\_type="thymocytes, peripheral blood lymphocytes"  
gene 1..3500  
1..3500  
/gene="JAK2"  
1..3399  
/gene="JAK2"  
/codon\_start=1  
/product="jak2 kinase"  
/protein\_id="AAC23653.1"  
/db\_xref="GI:3236322"

CDS

BASE COUNT 1181 a 608 c 740 g 971 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.86e-05 Length: 3500  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 9 Gaps: 0  
US-09-397-967-16 (1-1099) x AF001362 (1-3500)

QY 1003 SerAspValTrpSerPheGlyValValLeuTYRGluLeuPheThrTYR 1018  
DB 3103 TCAGATGTGTGGAGCTTTGGAGCTTCTATACGAACCTTTCACATAC 3150

RESULT 38

LOCUS AR136148 3629 bp DNA linear PAT 16-JUN-2001  
DEFINITION Sequence 8 from patent US 6136595.  
ACCESSION AR136148  
VERSION AR136148.1 GI:14476820  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3629)  
AUTHORS Ihle,J.F.Witthuhn,B.A. and Silvennoinen,O.  
TITLE Jak kinases and regulations of cytokine signal transduction  
JOURNAL Patent: US 6136595-A 8 24-OCT-2000;  
FEATURES Location/Qualifiers  
source 1..3629  
/organism="unknown"

BASE COUNT 1142 a 732 c 843 g 912 t  
ORIGIN

Alignment Scores:

Pred. No.: 5e-05 Length: 3629  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x ARI36148 (1-3629)

OY 1003 SeraspValtRrPserPheGlyValValLeuTyrgIuleupheThrTyr 1018  
 |||||||  
 DB 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGAACTTTTCACATAC 3243

RESULT 39

LOCUS ARI43945 3629 bp DNA linear PAT 08-AUG-2001  
 DEFINITION Sequence 8 from patent US 6210654.  
 ACCESSION ARI43945  
 VERSION ARI43945.1 GI:15105812  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 3629)

AUTHORS Ihle,J., Witthuhn,B.A., Quelle,F.W. and Silvennoinen,O.  
 TITLE Jak Kinases and Regulation of cytokine signal transduction  
 JOURNAL Patent: US 6210654-A 8 03-APR-2001;  
 FEATURES Location/Qualifiers  
 source 1..3629

BASE COUNT 1142 a 732 c 843 g 912 t  
 ORIGIN

Alignment Scores:

Pred. No.: 5e-05 Length: 3629  
 Score: 16.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x ARI43945 (1-3629)

OY 1003 SeraspValtRrPserPheGlyValValLeuTyrgIuleupheThrTyr 1018  
 |||||||  
 DB 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGAACTTTTCACATAC 3243

RESULT 40

LOCUS 192575 3629 bp DNA linear PAT 01-DEC-1998  
 DEFINITION Sequence 8 from patent US 5728536.  
 ACCESSION 192575  
 VERSION 192575.1 GI:3937045  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 3629)

AUTHORS Ihle,J.N., Witthuhn,B.A., Quelle,F.W. and Silvennoinen,O.  
 TITLE Jak Kinases and Regulation of cytokine signal transduction  
 JOURNAL Patent: US 5728536-A 8 17-MAR-1998;  
 FEATURES Location/Qualifiers  
 source 1..3629

BASE COUNT 1142 a 732 c 843 g 912 t  
 ORIGIN

Alignment Scores:

Pred. No.: 5e-05 Length: 3629  
 Score: 16.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x 192575 (1-3629)

OY 1003 SeraspValtRrPserPheGlyValValLeuTyrgIuleupheThrTyr 1018  
 |||||||  
 DB 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGAACTTTTCACATAC 3243

RESULT 41

MUSPTK 3629 bp mRNA linear ROD 27-JAN-1997  
 LOCUS Mouse protein-tyrosine kinase (JAK2) mRNA, complete cds.  
 DEFINITION L16956  
 ACCESSION L16956  
 VERSION L16956.1 GI:309463  
 KEYWORDS protein-tyrosine kinase.  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 3629)

AUTHORS Silvennoinen,O., Witthuhn,B.A., Quelle,F.W., Cleveland,J.L., Yi,T.  
 TITLE Structure of the murine Jak2 protein-tyrosine kinase and its role  
 JOURNAL in interleukin 3 signal transduction  
 MEDLINE Proc. Natl. Acad. Sci. U.S.A. 90 (18), 8429-8433 (1993)  
 PUBMED 8378315  
 FEATURES Location/Qualifiers  
 source 1..3629

gene

1..3629 /organism="Mus musculus"  
 1..3629 /db\_xref="taxon:10090"  
 94..3483 /gene="JAK2"  
 /gene="JAK2"  
 /codon\_start=1  
 /evidence=experimental  
 /product="protein-tyrosine kinase"  
 /protein\_id="AA041327.1"  
 /db\_xref="GI:409584"

CDS

/translation="MGACLTMTMETSTSPYHONGDIPGSANSVYKQIEPVLOYLY  
 HSLQAGEYLPKPSGEYVAEICVASKACGTPYVHNKFMALMSEIERIYTPNHYF  
 HIDESTRHDLIYRIREFPHMYCGSSRTYGVSRGAELPDLDDFWYSLFVQWRD  
 FVHGWLIVPVTHETOEELCMAYLDMMRIKKEKQDPLAVNSVYTFLLPKVRAKI  
 ODYHILTRKRIYRFRFIOFSQCKATANKLKYLIINLETLOSAREYTOFEVKEBA  
 RGPGEIEIPATIIITGNGGIGMSGRKESSETLEDOVLCPEPDIDVSIQAOANE  
 CSNRSRYVHAKODGKVLLETLELSIKRLFSVLDISYVLTDAHYLCKEAPPV  
 LENTHSNGHPISMDPAISIKKAKNGTGYLRCSKDKNKYETFAVRENVIEYK  
 HCLITKNEBGLNLTSTNRNLSLKLNLCTQMETVSDSITIFOTKCCPKRDKSN  
 LVFERTGISDVOISPTLORNNVNWVFKIRNEDLIENESLGQGTFTIKGVRE  
 VGDYGLAKTEFVILKVDKARHNSSEFPAASMSQLSKRLVLANGVCGEENLE  
 VOEFKFSIDTYLKKKNSINILMKLGVAKKOLAMAHFLEKSLINGVCARNILLI  
 REEDRRNGNPETIKLSDPGISITVLPDKILOERIIPVPPICEPKNLNATDKMSFG  
 TTIMETSCGCDKPLSALDSOKIOFYEDKQOLAPKWTETLANLINCMOYEDPRPAF  
 RAVIRDLNLSLTPDYELLETENDMLPNRIGALGFSCAFEDRDPFOPEERILKTLQOLG  
 KGNFGSVEMCRYPDLQNTGEVAVAKRLQSTHEHLDFERELEILSLQHDNIVTK  
 GVCYSAGRRNLRLIMEYLPYGLSLDYVLRKLERIDRKLLOYSQICQKMEYLGATKY  
 IHRDLATRNILIVENENRVKIGDFGLTAVLQDQKRYVKKRGESEPIFWAVQSLTESK  
 FSVASDWSGCVLYELFTYIEKSKSPVEVRMIGDKOGOMIVFHLIELLSNGRL  
 PRPGCCDEIIVITTECMNNVNSQRPFRDLSFQWIKGIV"

BASE COUNT 1142 a 732 c 842 g 913 t  
 ORIGIN

Alignment Scores:

Pred. No.: 5e-05 Length: 3629  
 Score: 16.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x MUSPTK (1-3629)

OY 1003 SeraspValtRrPserPheGlyValValLeuTyrgIuleupheThrTyr 1018  
 |||||||

DB 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGAACTTTTCACATAC 3243

RESULT 42  
RNU13396 3731 bp mRNA linear ROD 26-OCT-1995  
DEFINITION Rattus norvegicus protein-tyrosine kinase (JAK2) mRNA, complete cds.  
ACCESSION U13396  
VERSION U13396.1 GI:1040849  
KEYWORDS  
SOURCE rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 3731)  
AUTHORS Duhe, R.J., Rul, H., Greenwood, J.D., Garvey, K. and Farrar, W.L.  
TITLE Cloning of the gene encoding rat JAK2, a protein tyrosine kinase  
JOURNAL Gene 158 (2), 281-285 (1995)  
MEDLINE 95331631  
PUBMED 7607555  
REFERENCE 2 (bases 1 to 3731)  
AUTHORS Duhe, R.J.  
TITLE Direct Submission  
JOURNAL Submitted (12-AUG-1994) Roy J. Duhe, Program Resources, Inc./Dyncorp., National Cancer Institute - Frederick Cancer Research and Development Center, P.O. Box B, Frederick, MD 21702-1201, USA  
COMMENT On Oct 26, 1995 this sequence version replaced gi:595682.  
FEATURES  
source location/Qualifiers  
1..3731  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/cell\_line="Nb2-Sp"  
/cell\_type="pre-T lymphoma"  
/note="Long-Evans parent strain"  
1..3731  
/gene="JAK2"  
157..3555  
/gene="JAK2"  
/note="cytokine receptor-related"  
/codon\_start=1  
/product="protein tyrosine kinase"  
/protein\_id="AA79911.1"  
/db\_xref="GI:595683"  
/translation="MGMACTLMTMEGTSTSPAHNGDIPGANSYKQTEPVLYQVLY  
HSLQAGEDYLKFPNGEYAEELCYAASACGTPYHNMFLMSETERIMVPPNHVE  
HIDESTRNHILYRIRYFPHWYCSGNRTARYGSGAEPALIDPVMISYLAQMRHD  
FVHGMIKVPVETETOECECLGMAVLDMRIAKEKQDPTPLAVNSISYKTPLPQVRAKI  
QDYLITRRIRYRFRRIIOFSOCKATARNLKLYLINLTLLOSFAFYEOPEVESAE  
RQPSGEIEFATITITNGGIGSROGKHKESFTLEDOLYCDPPIIDIVTSKQAOE  
CSRSRYVYVHKODQKVEIEISLKEALSFSLIDGYRLTADAHVYCKEAPAV  
LENIHNCAGPISMDFAISRLKAGQDTLYLRCSPKDNKYFPLFAVERNVETEK  
HCLITNENGEYLSGTRKNSFLADLNLQCYOMTVRSDSIIFOFKCCPPPKRSN  
LIVFRTNGVSDVQLSPTLQRHNHVNQMVPHKIRNEDLINSELSGGTETKIFKGVARE  
VDDYGLHTEVLKLYLDAKHNYSSEFEASMSQSLSHKHLVLYNAGVCAQGENIL  
VOEFYFSGLDIYLLKKNKSNINILMKLGAQOLANAMHLEEKSLIHGVACAKNILLI  
REEDRTGNPPTIKLSDPGISTIVLPKOLIORIPVPRECELENKRLTATDKSPFG  
TTLMECSGDKPLSLDSQRLQLETDHQLPAPRWELANLINTCMYEDDEPRAP  
RAVIRDLNLSLTPDELLENLPMWRKALGSGAFEDRDPQFEERHKLFLQGLQ  
KNGFSGVEMCRDPLQDNTGEVAVAKLQHSSTEBHLDFEREIEILSKOHNIYAK  
GYCYSAGRNRLILMEYLPYGLSLDYLQKHKEIRIDHKLLOYTSQICKMEYLGTRK  
IHRDLATRNILVENENRKLIGDGLTKVLPQDEYKVKVPEGESPIEMWAPSLTESK  
FSVASDVMFSGVYIELFTYIEKRSRPSPEPRMNCIDNGOGMIYPHLIELLKNRRL  
PRPEGCDIYIYIMTECMNNNNVNSPSEFSLSLRDQILDSMAA"

BASE COUNT 1149 a 765 c 881 g 936 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5, 11e-05 Length: 3731  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0

DB: 10 Gaps: 0  
US-09-397-967-16 (1-1099) x RNU13396 (1-3731)  
Qy 1003 SeraspValTrrpSerPheGlyValLeuTyrCluLeuPheRhrTyr 1018  
Db 3259 TCAGATGTGTGACACTTTGGTGTGCTGTGTACGACACTTTTCACGAT 3306  
RESULT 43  
AF148993 3739 bp mRNA linear VRT 29-JUN-2000  
LOCUS AF148993  
DEFINITION Cyprinus carpio janus kinase 3 mRNA, complete cds.  
ACCESSION AF148993  
VERSION AF148993.1 GI:6690507  
KEYWORDS  
SOURCE Cyprinus carpio.  
ORGANISM Cyprinus carpio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprinus.  
REFERENCE 1 (bases 1 to 3739)  
AUTHORS Yin, Z. and Kwang, J.  
TITLE Molecular cloning of a carp Jak3 from activated leukocytes  
JOURNAL Vet. Immunol. Immunopathol. 74 (3-4), 225-236 (2000)  
MEDLINE 20264065  
PUBMED 10802290  
REFERENCE 2 (bases 1 to 3739)  
AUTHORS Yin, Z. and Kwang, J.  
TITLE Direct Submission  
JOURNAL Submitted (06-MAY-1999) Laboratory of Animal Health Biotechnology, Institute of Molecular Agrobiolgy, National University of Singapore, 1 Research Link, Singapore 117604, Singapore  
FEATURES  
source location/Qualifiers  
1..3739  
/organism="Cyprinus carpio"  
/strain="Koi"  
/db\_xref="taxon:7962"  
/cell\_type="head kidney leukocytes"  
/note="con A induced"  
49..3429  
/codon\_start=1  
/product="Janus kinase 3"  
/protein\_id="AAE24169.1"  
/db\_xref="GI:6690508"  
/translation="MSEEEVPLMKSERAGSOKSNMDSALQVLYLSPINSETPPTI  
STRHTAENVCAIAKASGILPVEFNLTALASQDLSVYVPPHVRKSEINIKVHRRV  
FEFSPPQGSASARYPSISRRICAVLDYCYTIDLPKQSSDPVSGGIGSPALSLQ  
OECLEMAVLDRMAKERNQSLGECNTTSTKSCUPETHRODIOIRSLARQIKRTL  
KFLKILKSCSVERILKLYLMELSVLEPNQSSFPPLHSGMLEQSEQORVKAVRY  
SGEGIIQITKESQEMQFCDFPQITDISIKRLSQEOMPLEQNVYTLTRDDQMEAE  
FHTLLEALSFWLSVGVFRLTDTDSYHSCAEVAPSLIEDIONYCHGPTSEFAVHL  
KKAAGKNMFLRLHSPKDFKDFVLCVCIOTHLGMYKCCLEKNKFLIAGIHNSFCS  
LKQLDPEFQHSILMSDIPVTLKCCPPRPRLTMMILIRNSSMEMSSPMLQHRKP  
SMQPHMILKHEDTBSLGGSTPHIFRGSTTDORQDVTTFEVLVLDANINCM  
BSFEASALMSQISIKHLIVGISVHRSKNIMVEFEVKAHALDLYLKKSYSVSMK  
LDVAQOLAKLSQIELEKNIHAGNICAKNLLAREDDPSTDSPTFKISDPIINNALIG  
KDVLADLIPMAVEVLDLLELEECDDKSGFTTLMELINNECAPLOGODLJHKOOFE  
HESNLPALREMLADLILAQCMQOELPSCRSITROLNSLITTSYETLHAAGTLPKS  
DGFMRPMKFOQOODEEREELRFTISLIGQNGFSEVLCRDPMDGNTGLVAVKEL  
OSNKATVTDQREIOTISLSLHCDIVYKGYCTGSLSKRWETIYPTSLGTYME  
KHROWCRRLLLPASQICKMEYIQSMRYHRRDAANRLVASDTLVKIDAGLTKI  
IPVDEYLRVTPGSEPVFWAPESISELKSFKSDIVSFGVILHELPSYDINSNPQ  
KIVPEDEYLRKCEPVPMWSIFLIFSKDMRLPAPQCCPPKHNIMMOCMEFNSEDRPS  
FSSLQDLIENSLDERECCKGIRVTLNRSLYFQIMYT"

BASE COUNT 1045 a 844 c 888 g 962 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5, 12e-05 Length: 3739  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0

DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x AFL48993 (1-3739)

QY 824 leuGLyLVsGLYAsnPhcGLySerValGluLeuGysArgTyAspPro 839  
|||||

Db 2527 CTTGGGAAGGACTCGGACAGTGTTCAGCTTTGTCTGTTAAGATCCG 2574

RESULT 44

AB036337 4042 bp mRNA linear MAM 08-JAN-2000

LOCUS Sus scrofa mRNA for kinase-defective JAK2 variant, complete cds.

ACCESSION AB036337.1 GI:6635253

VERSION AB036337.1

KEYWORDS kinase-defective JAK2 variant.

SOURCE Sus scrofa adult muscle thoracic aortic cell cDNA to mRNA, clone:50D1.

ORGANISM Sus scrofa

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

AUTHORS Ito, Y., Minezawa, M. and Awata, T.

TITLE Sus scrofa mRNA for JAK2 alternative splicing variant, complete cds

JOURNAL Published Only in Database (1999)

REFERENCE 2 (bases 1 to 4042)

AUTHORS Ito, Y., Minezawa, M. and Awata, T.

TITLE Direct Submission

AUTHORS Submitted (22-Dec-1999) Yoshiyasu Ito, National Institute of Animal Industry, Animal Genome Research Team, 2 Ikenodai, Kunitzaki-machi, Inshiki-gun, Ibaragi 305-0901, Japan

FEATURES

source 1. 4042

location/Qualifiers

1. 4042

/organism="Sus scrofa"

/db\_xref="taxon:9823"

/clone="50D1"

/cell\_type="thoracic aortic cell"

/tissue\_type="muscle"

/dev\_stage="adult"

1. 4042

/gene="JAK2"

407. .3697

/gene="JAK2"

/codon\_start=1

/product="kinase-defective JAK2 variant"

/protein\_id="BAA8632.1"

/db\_xref="GI:6635254"

/translation="MGMACLTPTMEGTSSTPYONGDIPGANSVKQIDPVLQVLYL  
HSLGAEGLYLFPSGEYVAEETIAASKAGCTTPTVTHNFMALSETRTWYPRNHV  
HYDESTRNVLRIRFPYPRYVNGSNRTYRGISRGAEAPLLDFVMSYLFPAQRHD  
FLYGWVKIPVTHETOEELGMAVLDMRIAKENQDTPLDIVSSVSKTEFLPCVRAKI  
ODYHLTRKRIYRFRRTIOFSGKATARNLKLKYLINLETLQSAFTEFEVKEPG  
RGSGEELFATIIITGNGGIOMSRGKHSEETLEODLOLCDEPNITIDVSIKOANOE  
GSNERVATITHODKSLIEILSLREALSVSLIDGYRLTAADAHYLCKEVAPRAY  
LENIQSNCHGPISMDFALSKLKAQNOGLVLRCSKDPFKYFLTRAVERNVIEYK  
HCLITKNENGEYVLSGTRKNFSLKDLNLCYOMETVRSNIIIPFTKCCPPKPKDSN  
LLVFRTNGISDVPTSLQRTAHNNVQMFHRI RNEDLI FNESLQGTFTPKIFKGVARE  
VGDYGLHETVELKLVLDKARNVSESEFEASMSKLSHKLVLVNYGVCAGDENIL  
VOEYKFSGLDYTLKKNKNSINILMKLEVAKOLAMAMHLEKTLIHGVACNILLI  
REEDRTKGNPPPTIKLSDPGISTTVLPKTILORIIPWVPECIENKNTLADTKMSFG  
TTLMETCSGGDKPLNALDSQRKLOFEDRHOLPAKRWELANLINCMADYEDPDRPSF  
RAIIRDLNLSLFTPDYELLTENDMLPNMRIGALGSGAEFDRDPTQFEERHUKFTLOQG  
KCNFSGVEMCRYPDLPDNTGGEVAVAKLQHGSTEELRDFERIEILSKQMDNIYVK  
GVCYAGRNRNLKIMEYLPYGLRDYLOKHKERIDHKIKLOYTSQICKMEYLGTRKY  
IHRDLATRNIIIVLENENRYKIGDGLTKVLPDKKEYKVPESPIFWAPESLIESK  
FSVADSWMSFGVYVLELFTYITKSKSPRAVSVLPNFQLEFFPNMSKAPPKIKLARI  
LEPHTTYL"

4042

/gene="JAK2"

BASE COUNT 1290 a 737 c 875 g 1140 t

ORIGIN

Alignment Scores:

Pred. No.: 5,45e-05 Length: 4042

Score: 16.00 Matches: 16

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.46% Indels: 0

DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x AB036337 (1-4042)

QY 1003 SeraspValTTPSerPhcGLYValValLeuTyrgluLeuPheThrTYR 1018  
|||||

Db 3509 TCACATGTTTGAGAGCTTGAGTGTCTGTATGTAACATTTACATAT 3556

RESULT 45

AF005216 4161 bp mRNA linear PRI 30-OCT-1997

LOCUS Homo sapiens receptor-associated tyrosine kinase (JAK2) mRNA,

DEFINITION complete cds.

ACCESSION AF005216

VERSION AF005216

KEYWORDS complete cds.

SOURCE AF005216.1 GI:2570358

ORGANISM Homo sapiens.

REFERENCE Homo sapiens.

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

TITLE 1 (bases 1 to 4161)

AUTHORS Peeters, P., Raynaud, S.D., Cools, J., Wlodarska, I., Grosgeorge, J., Philip, P., Monpoux, F., Van Rompaey, L., Baens, M., Van den Berghe, H. and Marynen, P.

TITLE Fusion of TEL, the ETS-variant gene 6 (ETV6), to the receptor-associated kinase JAK2 as a result of t(9;12) in a lymphoid and t(9;15;12) in a myeloid leukemia

JOURNAL Blood 90 (7), 2535-2540 (1997)

PUBMED 97465498

9326218

2 (bases 1 to 4161)

Peeters, P., Cools, J. and Marynen, P.

TITLE Direct Submission

AUTHORS Submitted (23-MAY-1997) Human Genome Laboratory, Center for Human Genetics, Herestraat 49, Leuven B-3000, Belgium

COMMENT On Oct 30, 1997 this sequence version replaced gi:2565219.

FEATURES

source 1. 4161

location/Qualifiers

1. 4161

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="9"

/map="9p24"

1. .4161

/gene="JAK2"

495. .3893

/gene="JAK2"

/note="Janus kinase 2"

/codon\_start=1

/product="receptor-associated tyrosine kinase"

/protein\_id="AAB82092.1"

/db\_xref="GI:2570359"

/translation="MGMACLTPTMEGTSSTPYONGDIPGANSVKQIDPVLQVLYL  
HSLGAEGLYLFPSGEYVAEETIAASKAGCTTPTVTHNFMALSETRTWYPRNHV  
HYDESTRNVLRIRFPYPRYVNGSNRTYRGISRGAEAPLLDFVMSYLFPAQRHD  
FLYGWVKIPVTHETOEELGMAVLDMRIAKENQDTPLDIVSSVSKTEFLPCVRAKI  
ODYHLTRKRIYRFRRTIOFSGKATARNLKLKYLINLETLQSAFTEFEVKEPG  
RGSGEELFATIIITGNGGIOMSRGKHSEETLEODLOLCDEPNITIDVSIKOANOE  
GSNERVATITHODKSLIEILSLREALSVSLIDGYRLTAADAHYLCKEVAPRAY  
LENIQSNCHGPISMDFALSKLKAQNOGLVLRCSKDPFKYFLTRAVERNVIEYK  
HCLITKNENGEYVLSGTRKNFSLKDLNLCYOMETVRSNIIIPFTKCCPPKPKDSN  
LLVFRTNGISDVPTSLQRTAHNNVQMFHRI RNEDLI FNESLQGTFTPKIFKGVARE  
VGDYGLHETVELKLVLDKARNVSESEFEASMSKLSHKLVLVNYGVCAGDENIL  
VOEYKFSGLDYTLKKNKNSINILMKLEVAKOLAMAMHLEKTLIHGVACNILLI  
REEDRTKGNPPPTIKLSDPGISTTVLPKTILORIIPWVPECIENKNTLADTKMSFG  
TTLMETCSGGDKPLNALDSQRKLOFEDRHOLPAKRWELANLINCMADYEDPDRPSF  
RAIIRDLNLSLFTPDYELLTENDMLPNMRIGALGSGAEFDRDPTQFEERHUKFTLOQG  
KCNFSGVEMCRYPDLPDNTGGEVAVAKLQHGSTEELRDFERIEILSKQMDNIYVK  
GVCYAGRNRNLKIMEYLPYGLRDYLOKHKERIDHKIKLOYTSQICKMEYLGTRKY



polyA\_site  
4599  
/gene="JAK2"  
BASE COUNT 1491 a 810 c 954 g 1344 t  
ORIGIN

Alignment Scores:  
Pred. No.: 6.04e-05 Length: 4599  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x AB036336 (1-4599)

QY 1003 SeraspValTrrpSerPheGlyValValleuTyrGluLeuPheThrTyr 1018  
|||||  
Db 3206 TCAGATGTTTGAGAGCTTTGGAGTGTCTGTATGACATTTTACACATAT 3253

RESULT 48  
AB006011 4812 bp mRNA linear MAM 29-JUL-1997  
LOCUS Sus scrofa mRNA for JAK2, complete cds.  
DEFINITION AB006011  
ACCESSION AB006011 GI:2285795  
VERSION JAK2.  
KEYWORDS Sus scrofa adult muscle thoracic aortic cell cDNA to mRNA.  
SOURCE Sus scrofa  
ORGANISM Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
REFERENCE 1 (sites)  
AUTHORS Ito, Y., Mikawa, S., Kobayashi, E., Wada, Y. and Minezawa, M.  
TITLE Domestic pig mRNA for JAK2, complete cds  
JOURNML Unpublished  
REFERENCE 2 (bases 1 to 4812)  
AUTHORS Ito, Y.  
TITLE Direct Submission  
JOURNML Submitted (24-JUL-1997) Yoshiyasu Ito, National Institute of Animal  
Industry, Animal Genome Research Team; 2 Ikenodai, Kikizaki-machi,  
Inashiki-gun, Ibaragi 305, Japan (E-mail: yoshiito@niai.affrc.go.jp,  
Tel:0298-38-8627)

FEATURES  
source location/Qualifiers  
1..4812  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/cell\_type="thoracic aortic cell"  
/tissue\_type="muscle"  
/dev\_stage="adult"  
317..3712  
/codon\_start=1  
/product="JAK2"  
/protein\_id="AA021662.1"  
/db\_xref="GI:2285796"  
/translation="MGMACTMTMEGTSTSPVHONGDIPGNANSVKOIDPVLOYLYL  
HSLGKAEDYLKFPAGEVVAEICIAASAKAGITPPYHNMFLMSETERIWPNNHVF  
HIDESTRHNVLYRIRFYPYWGSGSNRTYRHGISRGAEALDDFVMSYLFQWRHD  
FLYGVWKIPVTHETOEELGMAVLDMMRIAKEDOTPLDISSVSKFPLFCVAKI  
ODYHILTRKRIRYRFRRTIEOFSHCKATARNIKKLYLINLETLOSATFEKFEVKEPG  
RGPSGEELFATIIITNGGCIOMSRGKRRESFTLEODIOLXCPDSNIDVSIQANOE  
GSNSRVTIHKODKSLLEISLRALSRLSVSLDGYRRLTADAHYLCKEVAPAY  
LENIOSNCHGPIISMDFALSKLKAGNOTGLVLRCSFDENKVFILTFVARENVLEK  
HCLITKNENENYNSGTRKNFNSLKDILNCOYMETVRSDIIFQPTKCCPKRDKSN  
LLVPRINGVSDVPTSPLOPRTNHNQVFNHRIINREDLIFNESLGGTFTKIFGVRE  
VGDGOLAEHEFVILKLVLDKARNVSESEFEPAASMSKLSHKILVLYNGVCGEENIL  
VOEPYKSGSIDTYIKKKRKNSTIILMKLEVAOLAMAMHLEKTLIHGVCANILILI  
REEDRKGNPPIKLSDPGISTITVLPRILOERIPWPVPECTENKKNILATDKWSFG  
TTIMEICGGDKPLNALSOKRLOFEDRHOLPAKWAELANLNKMDYEDPRPSP  
RAIIRDNLSLTPPYELLTENDMLPNMRIGALGSGAFEDRDPQFEEHRLKFLQOLG  
KNGFSGVEMCRYPDLODNTGEVAVKIKLOHSTEEHRLDFEEREIELSLQHDNIVKRY  
GVCYSAGRNRKLIMEYLPYGLRDYLOKHKRIIDHKLLQYTSIQKGMELGTRKY  
IHRDLATRNILVENENRYKIGDFGLTKVLPQDKKEYKKKEGESPIFWVAPESLTSK

polyA\_site  
4812  
/note="17 a nucleotides"  
BASE COUNT 1554 a 853 c 1016 g 1389 t  
ORIGIN

Alignment Scores:  
Pred. No.: 6.27e-05 Length: 4812  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x AB006011 (1-4812)

QY 1003 SeraspValTrrpSerPheGlyValValleuTyrGluLeuPheThrTyr 1018  
|||||  
Db 3419 TCAGATGTTTGAGAGCTTTGGAGTGTCTGTATGACATTTTACACATAT 3466

RESULT 49  
AX354568 5117 bp DNA linear PAT 06-FEB-2002  
LOCUS Sequence 15 from Patent WO0179555.  
DEFINITION AX354568  
ACCESSION AX354568  
VERSION AX354568.1 GI:18619370  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Hancock, M.W. and Ozkaynak, E.  
TITLE Roles of Jak/stat family members in tolerance induction  
JOURNML Patent: WO 0179555-A 15 25-OCT-2001;  
Millennium Pharmaceuticals, Inc. (US)

FEATURES  
source location/Qualifiers  
1..5117  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/db\_xref="GI:18619371"  
/translation="MGMACTMTMEGTSTSSIYQNDISGNANSKKOIDPVLOYLYL  
HSLGKAEDYLKFPAGEVVAEICIAASAKAGITPPYHNMFLMSETERIWPNNHVF  
HIDESTRHNVLYRIRFYPYWGSGSNRTYRHGISRGAEALDDFVMSYLFQWRHD  
FLYGVWKIPVTHETOEELGMAVLDMMRIAKEDOTPLDISSVSKFPLFCVAKI  
ODYHILTRKRIRYRFRRTIEOFSHCKATARNIKKLYLINLETLOSATFEKFEVKEPG  
RGPSGEELFATIIITNGGCIOMSRGKRRESFTLEODIOLXCPDSNIDVSIQANOE  
GSNSRVTIHKODKSLLEISLRALSRLSVSLDGYRRLTADAHYLCKEVAPAY  
LENIOSNCHGPIISMDFALSKLKAGNOTGLVLRCSFDENKVFILTFVARENVLEK  
HCLITKNENENYNSGTRKNFNSLKDILNCOYMETVRSDIIFQPTKCCPKRDKSN  
LLVPRINGVSDVPTSPLOPRTNHNQVFNHRIINREDLIFNESLGGTFTKIFGVRE  
VGDGOLAEHEFVILKLVLDKARNVSESEFEPAASMSKLSHKILVLYNGVCGEENIL  
VOEPYKSGSIDTYIKKKRKNSTIILMKLEVAOLAMAMHLEKTLIHGVCANILILI  
REEDRKGNPPIKLSDPGISTITVLPRILOERIPWPVPECTENKKNILATDKWSFG  
TTIMEICGGDKPLNALSOKRLOFEDRHOLPAKWAELANLNKMDYEDPRPSP  
RAIIRDNLSLTPPYELLTENDMLPNMRIGALGSGAFEDRDPQFEEHRLKFLQOLG  
KNGFSGVEMCRYPDLODNTGEVAVKIKLOHSTEEHRLDFEEREIELSLQHDNIVKRY  
GVCYSAGRNRKLIMEYLPYGLRDYLOKHKRIIDHKLLQYTSIQKGMELGTRKY  
IHRDLATRNILVENENRYKIGDFGLTKVLPQDKKEYKKKEGESPIFWVAPESLTSK



Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AX354568 (1-5117)

OY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
 DB 3580 TCAGATGTTTGAGCTTGAGCTGTTCTGTATGAACCTTTCACATAC 3627

RESULT 50  
 AF058925 5117 bp mRNA linear PRI 23-JUN-1998  
 LOCUS Homo sapiens Jak2 kinase mRNA, complete cds.  
 DEFINITION AF058925  
 ACCESSION AF058925  
 VERSION AF058925.1 GI:3068751  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 Saltzman, A., Stone, M., Franks, C., Searfoss, G., Munro, R., Jaye, M.,  
 and Ivashchenko, Y.  
 Cloning and characterization of human Jak-2 kinase: high mRNA  
 expression in immune cells and muscle tissue  
 Blochem. Biophys. Res. Commun. 246 (3), 627-633 (1998)  
 9618263

JOURNAL MEDLINE  
 PUBMED 9618263

REFERENCE 2 (bases 1 to 5117)  
 Saltzman, A., Stone, M., Franks, C., Searfoss, G., Munro, R., Jaye, M.,  
 and Ivashchenko, Y.  
 Direct Submission  
 Submitted (10-APR-1998) Gene Medicine, Rhone-Poulenc Rorer, 500  
 Arcola Rd., Collegeville, PA 19426, USA

TITLE JOURNAL  
 FEATURES Location/Qualifiers  
 source 1..5117  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1..2393  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="I.M.A.G.E Consortium ID 686684"  
 478..3876  
 /codon\_start=1  
 /product="Jak2 kinase"  
 /protein\_id="AAC23982.1"  
 /translation="MGACLTMTMEGNTSSITONGDISGNANSMGIDPVALOYL  
 HSLKSPDYITFPSEGEVAEITIASKACGTPVYHNMFALMSETERIVPNNVF  
 HIDESTRHNVLRIREFPRMYCGSNKAYRHGISRGAEPLDIDFVSLFAQRHD  
 FVHGIVPVYTHETQEDCLGAVLDMMHIAENDQTPLATINSISYTFELKCRAKI  
 ODVHILTRKIRIRYRFRFIOFSQCKATARKLKLYLINTLEDSAFTEFEVEREG  
 SGSPGEIIFATIIITNGNGIOWSRGKRESELTLEODOLICDSDSNIIIDVIRKANE  
 GSNESRYVTHIKODCKNLEIETLSRELSVSLIDCYRYLTAAHHYLCIEVAPRY  
 LENTOSNCHGISMDPAISKKAGNOTGLVLCSPKDFKITLTRAVERNENYETK  
 HCLITKNENEYVNSGTRKNESLKLDLNLCYOMETVSDNIIIFQTCPCPRDKSN  
 LIAVPTNGVSDVPTSPTRQPTNNQVFNHRIENEDLIFNSLOQGTFTKFKVIRE  
 VGDYQGLHETEVLLVLDKAKHNRSESEFEASMSKLSHKLVLNLCVCGENIL  
 VOEFVFGSLDTYLLKKNKNCINILMKLEVAOKLAWAMHLEENLLIGNVAKNILLI  
 REEDRTGNPPEIKTSDPGISITVLPKDIIOERTPWPRECIENPKNLNLTDMWSR  
 TTLWEISGCDKPLSALDSOKLQPTEDRHOLPAKMAELANLINNCDYEDPDRPSF  
 RAILRLDNLSTLPDYELLTENDMLPNMRIGALGSGAFEDDPQFEBRHLKLOQL  
 KGNFSGVEMCRYPDLQNTGEVAVAKLQHSTEBHLRDEFEIEILKSLQDNIVKRY  
 GVCYSAGRNLKILMEYLPYGLSDYLOKHREIDHILKLOYTQICKMGYLTGTRY  
 IHRDLATRNILVENENVKIGDFGLTKVLPQDKERYVKEGESPIRMVYAPESLTEK  
 FSVASDWSRGCVLVYELFTYTERKSPAPERMICNDKOCQIVPHILILKNNRL  
 PRPDCDEIIMTEECMNNVNRPSFRDLAKVDOIRDMAG"

BASE COUNT 1633 a 927 c 1103 g 1464 t

ORIGIN

Alignment Scores: 6.58e-05 Length: 5117

Score: 16.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AF058925 (1-5117)

OY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
 DB 3580 TCAGATGTTTGAGCTTGAGCTGTTCTGTATGAACCTTTCACATAC 3627

RESULT 51  
 AC119228/6  
 LOCUS Mus musculus clone RP24-209J17, LOW-PASS SEQUENCE SAMPLING.  
 DEFINITION AC119228 66974 bp DNA linear HTG 25-APR-2002  
 AC119228.1 GI:20303952  
 ACCESSION AC119228  
 VERSION AC119228.1 GI:20303952  
 KEYWORDS HTG: HTGS\_PHA50.  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Birren, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N.,  
 Anderson, S., Barina, N., Baslet, V., Bloom, T., Boguslavsky, L.,  
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,  
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collimore, A.,  
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
 Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Glinde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,  
 Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
 Maclean, C., Macdonald, P., Major, J., Marcus, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,  
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
 Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,  
 Zainoun, J., Zembek, L., Zlimer, A. and Zody, M.  
 Direct Submission  
 Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 smt, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE JOURNAL  
 COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIRB  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L25230  
 Center clone name: 209\_J\_17

NOTE: This record contains 83 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone

\* Will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved. 709: contig of 709 bp in length  
\* 1  
\* 710 809: gap of 100 bp  
\* 810 1526: contig of 717 bp in length  
\* 1527 1626: gap of 100 bp  
\* 1627 2342: contig of 716 bp in length  
\* 2343 2442: gap of 100 bp  
\* 2443 3133: contig of 691 bp in length  
\* 3134 3233: gap of 100 bp  
\* 3234 4033: contig of 700 bp in length  
\* 3934 4033: gap of 100 bp  
\* 4034 4730: contig of 697 bp in length  
\* 4731 4830: gap of 100 bp  
\* 4831 5519: contig of 689 bp in length  
\* 5520 5619: gap of 100 bp  
\* 5620 6322: contig of 703 bp in length  
\* 6323 6422: gap of 100 bp  
\* 6423 7140: contig of 718 bp in length  
\* 7141 7240: gap of 100 bp  
\* 7241 7964: contig of 724 bp in length  
\* 7965 8064: gap of 100 bp  
\* 8065 8793: contig of 729 bp in length  
\* 8794 8893: gap of 100 bp  
\* 8894 9622: contig of 729 bp in length  
\* 9623 9722: gap of 100 bp  
\* 9723 10458: contig of 736 bp in length  
\* 10459 10558: gap of 100 bp  
\* 10560 11247: contig of 689 bp in length  
\* 11248 11347: gap of 100 bp  
\* 11348 12047: contig of 700 bp in length  
\* 12048 12147: gap of 100 bp  
\* 12149 12842: contig of 695 bp in length  
\* 12843 12942: gap of 100 bp  
\* 12943 13635: contig of 694 bp in length  
\* 13637 13735: gap of 100 bp  
\* 13737 14466: contig of 730 bp in length  
\* 14467 14566: gap of 100 bp  
\* 14567 15276: contig of 710 bp in length  
\* 15277 15376: gap of 100 bp  
\* 15377 16098: contig of 722 bp in length  
\* 16099 16198: gap of 100 bp  
\* 16199 16924: contig of 726 bp in length  
\* 16925 17024: gap of 100 bp  
\* 17025 17760: contig of 736 bp in length  
\* 17761 17860: gap of 100 bp  
\* 17861 18550: contig of 690 bp in length  
\* 18551 18650: gap of 100 bp  
\* 18651 19353: contig of 703 bp in length  
\* 19354 19453: gap of 100 bp  
\* 19454 20158: contig of 705 bp in length  
\* 20159 20258: gap of 100 bp  
\* 20259 20973: contig of 715 bp in length  
\* 20974 21073: gap of 100 bp  
\* 21074 21780: contig of 707 bp in length  
\* 21781 21880: gap of 100 bp  
\* 21881 22600: contig of 720 bp in length  
\* 22601 22700: gap of 100 bp  
\* 22701 23397: contig of 697 bp in length  
\* 23398 23497: gap of 100 bp  
\* 23498 24216: contig of 719 bp in length  
\* 24217 24316: gap of 100 bp  
\* 24317 25052: contig of 736 bp in length  
\* 25053 25152: gap of 100 bp  
\* 25153 25888: contig of 736 bp in length  
\* 25889 25988: gap of 100 bp  
\* 25989 26692: contig of 704 bp in length  
\* 26693 26792: gap of 100 bp  
\* 26793 27478: contig of 686 bp in length  
\* 27479 27578: gap of 100 bp  
\* 27579 28290: contig of 712 bp in length  
\* 28291 28390: gap of 100 bp

\* 28391 29098: contig of 708 bp in length  
\* 29099 29198: gap of 100 bp  
\* 29199 29903: contig of 705 bp in length  
\* 29904 30003: gap of 100 bp  
\* 30004 30709: contig of 706 bp in length  
\* 30710 30809: gap of 100 bp  
\* 30810 31535: contig of 726 bp in length  
\* 31536 31635: gap of 100 bp  
\* 31636 32329: contig of 694 bp in length  
\* 32330 32429: gap of 100 bp  
\* 32430 33137: contig of 708 bp in length  
\* 33138 33237: gap of 100 bp  
\* 33238 33969: contig of 732 bp in length  
\* 33970 34069: gap of 100 bp  
\* 34070 34760: contig of 691 bp in length  
\* 34761 34860: gap of 100 bp  
\* 34861 35566: contig of 706 bp in length  
\* 35567 35666: gap of 100 bp  
\* 35667 36371: contig of 705 bp in length  
\* 36372 36471: gap of 100 bp  
\* 36472 37166: contig of 695 bp in length  
\* 37167 37266: gap of 100 bp  
\* 37267 37982: contig of 716 bp in length  
\* 37983 38082: gap of 100 bp  
\* 38083 38784: contig of 702 bp in length  
\* 38785 38884: gap of 100 bp  
\* 38885 39591: contig of 707 bp in length  
\* 39592 39691: gap of 100 bp  
\* 39692 40413: contig of 722 bp in length  
\* 40414 40513: gap of 100 bp  
\* 40514 41246: contig of 733 bp in length  
\* 41247 41346: gap of 100 bp  
\* 41347 42078: contig of 732 bp in length  
\* 42079 42178: gap of 100 bp  
\* 42179 42874: contig of 696 bp in length  
\* 42875 42974: gap of 100 bp  
\* 42975 43677: contig of 703 bp in length  
\* 43678 43777: gap of 100 bp  
\* 43778 44472: contig of 695 bp in length  
\* 44473 44572: gap of 100 bp  
\* 44573 45279: contig of 707 bp in length  
\* 45280 45379: gap of 100 bp  
\* 45380 46053: contig of 674 bp in length  
\* 46054 46153: gap of 100 bp  
\* 46154 46860: contig of 707 bp in length  
\* 46861 46960: gap of 100 bp  
\* 46961 47646: contig of 686 bp in length  
\* 47647 47746: gap of 100 bp  
\* 47747 48458: contig of 712 bp in length  
\* 48459 48558: gap of 100 bp  
\* 48559 49262: contig of 704 bp in length  
\* 49263 49362: gap of 100 bp  
\* 49363 50090: contig of 728 bp in length  
\* 50091 50190: gap of 100 bp  
\* 50191 50902: contig of 712 bp in length  
\* 50903 51002: gap of 100 bp  
\* 51003 51740: contig of 738 bp in length  
\* 51741 51840: gap of 100 bp  
\* 51841 52537: contig of 697 bp in length  
\* 52538 52637: gap of 100 bp  
\* 52638 53332: contig of 695 bp in length  
\* 53333 53432: gap of 100 bp  
\* 53433 54094: contig of 662 bp in length  
\* 54095 54194: gap of 100 bp  
\* 54195 54877: contig of 683 bp in length  
\* 54878 54977: gap of 100 bp  
\* 54978 55686: contig of 709 bp in length  
\* 55687 55786: gap of 100 bp

Alignment Scores:  
Pred. No.: 0.000515 Length: 66974  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x AC119228 (1-66974)

OY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
 Db 58447 TCAGATGTGTGAGCTTGTGAGTGTCTATACGAACTTTTCACATAC 58400

RESULT 52  
 AC096324\_3/c  
 WPCOMMENT

Sequence split into 4 fragments LOCUS AC096324 Accession AC096324  
 Fragment Name Begin End  
 AC096324\_0 1 110000  
 AC096324\_1 100001 210000  
 AC096324\_2 200001 310000  
 AC096324\_3 300001 385671

Continuation (4 of 4) of AC096324 from base 300001 (AC096324 Rattus norvegicus clone CH2

Alignment Scores:  
 Pred. No.: 0.00627 length: 85671  
 Score: 16.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x AC096324\_3 (1-85671)

OY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
 Db 80889 TCAGATGTGTGAGCTTGTGAGTGTCTATACGAACTTTTCACATAC 80842

RESULT 53  
 AL161450  
 LOCUS

DEFINITION Human DNA sequence from clone RPl1-39K24 on chromosome 9 Contains the JAK2 (Janus kinase 2 (a protein tyrosine kinase)) gene, a pseudogene similar to poly(ADP-ribose) synthetase, 2 pseudogenes similar to NADH-ubiquinone/plastoquinone, a pseudogene similar to cytochrome C oxidase, a pseudogene similar to TCF3 (transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)), IGHEP2 (immunoglobulin epsilon pseudogene 2), the INS16 (insulin-like 6) gene and Cpg islands, complete sequence.

ACCESSION AL161450 171146 bp DNA linear PRI 04-DEC-2001  
 VERSION AL161450.14 GI:9944145  
 KEYWORDS HTG: Cpg Island; Cytochrome; IGHEP2; Immunoglobulin; INS16; Insulin; JAK2; Janus; Kinase; NADH; oxidase; plastoquinone; poly(ADP-ribose) synthetase; TCF3; tyrosine; Ubiquinone; human.

SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 171146)  
 TITLE Direct Submission  
 AUTHORS Submitted (04-DEC-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humuery@sanger.ac.uk  
 JOURNAL On Aug 29, 2000 this sequence version replaced gi:9856709.  
 COMMENT During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission only a small overlap as described above.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence

FEATURES  
 source  
 misc-feature  
 gene  
 mRNA  
 CDS

was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9>  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RPl1-39K24 is from the library RPl1-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACE3.6  
 IMPORTANT: This sequence is not the entire insert of clone RPl1-39K24. It may be shorter because we sequence overlapping sections only once, except for a short overlap.  
 The true left end of clone RPl1-39K24 is at 1 in this sequence. The true left end of clone RPl1-140C18 is at 171047 in this sequence.

Location/Qualifiers  
 1..171146  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="9"  
 /clone="RPl1-39K24"  
 /clone\_1lb="RPl1-11.1"  
 /complement(3009..3495)  
 /note="match: GSS: Em:AO223474"  
 join(3254..3504,11074..11197,25694..25811,31977..32122,35854..36175,36960..37079,46174..46331,47969..48080,50313..50499,51216..51343,53783..53917,54989..55076,58744..58871,59597..59735,61520..61671,61824..61974,63016..63152,70965..71154,71737..71861,72030..72202,104295..104412,107624..107737,107975..109474)  
 /gene="JAK2"  
 join(3254..3504,11074..11197,25694..25811,31977..32122,35854..36175,36960..37079,46174..46331,47969..48080,50313..50499,51216..51343,53783..53917,54989..55076,58744..58871,59597..59735,61520..61671,61824..61974,63016..63152,70965..71154,71737..71861,72030..72202,104295..104412,107624..107737,107975..109474)  
 /gene="JAK2"  
 /product="pA39K24.1 (Janus kinase 2 (a protein tyrosine kinase))"  
 /note="match: CDNAS: Em:U13396 Em:AF001362 Em:AF058925 Em:U16956 Em:AB006011 Em:AF005216 Em:AB036336 Em:AB036337 Em:AB005690 Em:AB0000557  
 match: ESTs: Em:AA259122 Em:AA521703 Em:AM604114"  
 /evidence="not\_experimental"  
 join(3279..3504,11074..11197,25694..25811,31977..32122,35854..36175,36960..37079,46174..46331,47969..48080,50313..50499,51216..51343,53783..53917,54989..55076,58744..58871,59597..59735,61520..61671,61824..61974,63016..63152,70965..71154,71737..71861,72030..72202,104295..104412,107624..107737,107975..109474)  
 /gene="JAK2"  
 /note="match: proteins: Tr:O93595 Tr:O9PV12 Tr:O42291 Tr:O19064 Sw:O60674 Sw:O63272 Sw:O62689 Sw:O62120 Tr:O9Y652 Tr:O9YTU0 Tr:O35804"  
 /codon\_start=1  
 /evidence="not\_experimental"  
 /product="pA39K24.1 (Janus kinase 2 (a protein tyrosine kinase))"  
 /protein\_id="CAD13329.1"  
 /db\_xref="GI:17384438"  
 /translation="MGACLTWTEEGTSSITONDISGANSKMOIDPVLGVLYHSLKSEADYLFPSGEVVAEICIIASKAGITPVVHNMFLSETERIVPNVYFHIIDSTRNIVYLRFRFPRVYSGSNAYAGISRAEALDDPVMVSYFAPWRD FVHMIVKPVTHREOEELGMAVLDMMRIAKENQOTPLATNSISYTPLPKICRAKI ODVHILFKRIRIRFRRRIIOFSCAKTARVLKIKYLINLETLOSATVEKFEKPG SCGPGEFAFIITNGCIOWSRGKRESTRITLEDOLQICDPPNITISIKANDE GSNSRVYTHIKDGKNIETLSSLRNLSFVSLIDGYRFLTADAHYLCKEVAPAV

LENIQNSGKHSIMDFAIKIKKAKGQGLVYLRCSPKQFNKYFPLTEFAVERNVE  
HCLITIRNENHEVLTSTKTLQKNCNINLMKLEVAKOLAMAMHLENTLIGHVCARNL  
LVLFRTNGSDVFTSPTLQRPTRHMQMTHKTRINEDLTINESLGGCTTKIKRGR  
VGGYQJLHETVELLVKLVKADAHRYSESEFEASMSKUSHKHLYLNTGVCQGDEN  
VOGEFVGSJDTLTKKNCININLMKLEVAKOLAMAMHLENTLIGHVCARNL  
REEDRTGNGPFIKLSDDPISITVLEPKEDILQERIPWAPPECIEENKRNINLADKMS  
TTLMETCSGDDKTLSDOSORLOFEDEDRHOPLAPKMAELANLINCKNDIPEDFPR  
RALIIROLNLSPPDVELLTENMDLPMYRIGALFSGAEFDRPPOFEEHKLFLQO  
KKNFGSYEMKRPPLDNTNGEVAYAKKLOHSHSEHJRDPEREIELTKSLQHNLYAK  
GVCYSAGRNKLKLTMETLPPGSLRDYLQHKERIDHITKILQTSQCKMELTGRK  
IHHDLATRNLIYENNRVAKIGDFGLTKVLPQKEIKYKAPESSTFWAPSLTE  
FVSASVDMSGVALVELTETFIKESKSPAEFMRKMGINDQGMIVFHLIELLKNGG  
PRDDGCDPEIYMTCTGMNNVNVORSPFDLALRVQDLIDNMAG"  
4307..4642  
repeat\_region  
/note="H1MA4A repeat: matches 5811..6155 of consensus"  
/note=match: GSS: Em:AQ128841"  
misc\_feature  
/note="match: GSS: Em:AQ128841"  
repeat\_region  
6716..6807  
/note="46 copies 2 mer tt 66% conserved"  
7989..8028  
/note="20 copies 2 mer tg 80% conserved"  
repeat\_region  
8104..8313  
/note="Charliez repeat: matches 3530..3762 of consensus"  
repeat\_region  
8499..10889  
/note="TIGER1 repeat: matches 1..2416 of consensus"  
misc\_feature  
complement(12366..12931)  
/note="match: GSS: Em:AQ417363"  
22026..23724  
/gene="JAK2"  
misc\_feature  
/note="Cpg island"  
gene  
/evidence=not\_experimental  
22236..22321  
/gene="ba39K24.2"  
CDS  
/pseudo  
22232..22321  
/gene="ba39K24.2"  
/note="ba39K24.2 (casein kinase pseudogene)  
match: proteins: Tr:Q9YEX2 Sw:Q62761 Sw:Q62762 Sw:Q62763 Sw:Q62764 Sw:Q62765 Sw:Q62766 Sw:Q62767 Sw:Q62768 Sw:Q62769 Sw:Q62770 Sw:Q62771 Sw:Q62772 Sw:Q62773 Sw:Q62774 Sw:Q62775 Sw:Q62776 Sw:Q62777 Sw:Q62778 Sw:Q62779 Sw:Q62780 Sw:Q62781 Sw:Q62782 Sw:Q62783 Sw:Q62784 Sw:Q62785 Sw:Q62786 Sw:Q62787 Sw:Q62788 Sw:Q62789 Sw:Q62790 Sw:Q62791 Sw:Q62792 Sw:Q62793 Sw:Q62794 Sw:Q62795 Sw:Q62796 Sw:Q62797 Sw:Q62798 Sw:Q62799 Sw:Q62800 Sw:Q62801 Sw:Q62802 Sw:Q62803 Sw:Q62804 Sw:Q62805 Sw:Q62806 Sw:Q62807 Sw:Q62808 Sw:Q62809 Sw:Q62810 Sw:Q62811 Sw:Q62812 Sw:Q62813 Sw:Q62814 Sw:Q62815 Sw:Q62816 Sw:Q62817 Sw:Q62818 Sw:Q62819 Sw:Q62820 Sw:Q62821 Sw:Q62822 Sw:Q62823 Sw:Q62824 Sw:Q62825 Sw:Q62826 Sw:Q62827 Sw:Q62828 Sw:Q62829 Sw:Q62830 Sw:Q62831 Sw:Q62832 Sw:Q62833 Sw:Q62834 Sw:Q62835 Sw:Q62836 Sw:Q62837 Sw:Q62838 Sw:Q62839 Sw:Q62840 Sw:Q62841 Sw:Q62842 Sw:Q62843 Sw:Q62844 Sw:Q62845 Sw:Q62846 Sw:Q62847 Sw:Q62848 Sw:Q62849 Sw:Q62850 Sw:Q62851 Sw:Q62852 Sw:Q62853 Sw:Q62854 Sw:Q62855 Sw:Q62856 Sw:Q62857 Sw:Q62858 Sw:Q62859 Sw:Q62860 Sw:Q62861 Sw:Q62862 Sw:Q62863 Sw:Q62864 Sw:Q62865 Sw:Q62866 Sw:Q62867 Sw:Q62868 Sw:Q62869 Sw:Q62870 Sw:Q62871 Sw:Q62872 Sw:Q62873 Sw:Q62874 Sw:Q62875 Sw:Q62876 Sw:Q62877 Sw:Q62878 Sw:Q62879 Sw:Q62880 Sw:Q62881 Sw:Q62882 Sw:Q62883 Sw:Q62884 Sw:Q62885 Sw:Q62886 Sw:Q62887 Sw:Q62888 Sw:Q62889 Sw:Q62890 Sw:Q62891 Sw:Q62892 Sw:Q62893 Sw:Q62894 Sw:Q62895 Sw:Q62896 Sw:Q62897 Sw:Q62898 Sw:Q62899 Sw:Q62900 Sw:Q62901 Sw:Q62902 Sw:Q62903 Sw:Q62904 Sw:Q62905 Sw:Q62906 Sw:Q62907 Sw:Q62908 Sw:Q62909 Sw:Q62910 Sw:Q62911 Sw:Q62912 Sw:Q62913 Sw:Q62914 Sw:Q62915 Sw:Q62916 Sw:Q62917 Sw:Q62918 Sw:Q62919 Sw:Q62920 Sw:Q62921 Sw:Q62922 Sw:Q62923 Sw:Q62924 Sw:Q62925 Sw:Q62926 Sw:Q62927 Sw:Q62928 Sw:Q62929 Sw:Q62930 Sw:Q62931 Sw:Q62932 Sw:Q62933 Sw:Q62934 Sw:Q62935 Sw:Q62936 Sw:Q62937 Sw:Q62938 Sw:Q62939 Sw:Q62940 Sw:Q62941 Sw:Q62942 Sw:Q62943 Sw:Q62944 Sw:Q62945 Sw:Q62946 Sw:Q62947 Sw:Q62948 Sw:Q62949 Sw:Q62950 Sw:Q62951 Sw:Q62952 Sw:Q62953 Sw:Q62954 Sw:Q62955 Sw:Q62956 Sw:Q62957 Sw:Q62958 Sw:Q62959 Sw:Q62960 Sw:Q62961 Sw:Q62962 Sw:Q62963 Sw:Q62964 Sw:Q62965 Sw:Q62966 Sw:Q62967 Sw:Q62968 Sw:Q62969 Sw:Q62970 Sw:Q62971 Sw:Q62972 Sw:Q62973 Sw:Q62974 Sw:Q62975 Sw:Q62976 Sw:Q62977 Sw:Q62978 Sw:Q62979 Sw:Q62980 Sw:Q62981 Sw:Q62982 Sw:Q62983 Sw:Q62984 Sw:Q62985 Sw:Q62986 Sw:Q62987 Sw:Q62988 Sw:Q62989 Sw:Q62990 Sw:Q62991 Sw:Q62992 Sw:Q62993 Sw:Q62994 Sw:Q62995 Sw:Q62996 Sw:Q62997 Sw:Q62998 Sw:Q62999 Sw:Q63000 Sw:Q63001 Sw:Q63002 Sw:Q63003 Sw:Q63004 Sw:Q63005 Sw:Q63006 Sw:Q63007 Sw:Q63008 Sw:Q63009 Sw:Q63010 Sw:Q63011 Sw:Q63012 Sw:Q63013 Sw:Q63014 Sw:Q63015 Sw:Q63016 Sw:Q63017 Sw:Q63018 Sw:Q63019 Sw:Q63020 Sw:Q63021 Sw:Q63022 Sw:Q63023 Sw:Q63024 Sw:Q63025 Sw:Q63026 Sw:Q63027 Sw:Q63028 Sw:Q63029 Sw:Q63030 Sw:Q63031 Sw:Q63032 Sw:Q63033 Sw:Q63034 Sw:Q63035 Sw:Q63036 Sw:Q63037 Sw:Q63038 Sw:Q63039 Sw:Q63040 Sw:Q63041 Sw:Q63042 Sw:Q63043 Sw:Q63044 Sw:Q63045 Sw:Q63046 Sw:Q63047 Sw:Q63048 Sw:Q63049 Sw:Q63050 Sw:Q63051 Sw:Q63052 Sw:Q63053 Sw:Q63054 Sw:Q63055 Sw:Q63056 Sw:Q63057 Sw:Q63058 Sw:Q63059 Sw:Q63060 Sw:Q63061 Sw:Q63062 Sw:Q63063 Sw:Q63064 Sw:Q63065 Sw:Q63066 Sw:Q63067 Sw:Q63068 Sw:Q63069 Sw:Q63070 Sw:Q63071 Sw:Q63072 Sw:Q63073 Sw:Q63074 Sw:Q63075 Sw:Q63076 Sw:Q63077 Sw:Q63078 Sw:Q63079 Sw:Q63080 Sw:Q63081 Sw:Q63082 Sw:Q63083 Sw:Q63084 Sw:Q63085 Sw:Q63086 Sw:Q63087 Sw:Q63088 Sw:Q63089 Sw:Q63090 Sw:Q63091 Sw:Q63092 Sw:Q63093 Sw:Q63094 Sw:Q63095 Sw:Q63096 Sw:Q63097 Sw:Q63098 Sw:Q63099 Sw:Q63100 Sw:Q63101 Sw:Q63102 Sw:Q63103 Sw:Q63104 Sw:Q63105 Sw:Q63106 Sw:Q63107 Sw:Q63108 Sw:Q63109 Sw:Q63110 Sw:Q63111 Sw:Q63112 Sw:Q63113 Sw:Q63114 Sw:Q63115 Sw:Q63116 Sw:Q63117 Sw:Q63118 Sw:Q63119 Sw:Q63120 Sw:Q63121 Sw:Q63122 Sw:Q63123 Sw:Q63124 Sw:Q63125 Sw:Q63126 Sw:Q63127 Sw:Q63128 Sw:Q63129 Sw:Q63130 Sw:Q63131 Sw:Q63132 Sw:Q63133 Sw:Q63134 Sw:Q63135 Sw:Q63136 Sw:Q63137 Sw:Q63138 Sw:Q63139 Sw:Q63140 Sw:Q63141 Sw:Q63142 Sw:Q63143 Sw:Q63144 Sw:Q63145 Sw:Q63146 Sw:Q63147 Sw:Q63148 Sw:Q63149 Sw:Q63150 Sw:Q63151 Sw:Q63152 Sw:Q63153 Sw:Q63154 Sw:Q63155 Sw:Q63156 Sw:Q63157 Sw:Q63158 Sw:Q63159 Sw:Q63160 Sw:Q63161 Sw:Q63162 Sw:Q63163 Sw:Q63164 Sw:Q63165 Sw:Q63166 Sw:Q63167 Sw:Q631

FEATURES	source	gene	CDS
misc_feature		complement(51942..52378)	
repeat_region		/note="match: GSS: Em:A0230025"	
misc_feature		51995..52071	
repeat_region		/note="trigger2 repeat: matches 1..77 of consensus"	
misc_feature		52558..52871	
repeat_region		/gene="JAK2"	
repeat_region		/note="match: GSS: Em:A0074500"	
misc_feature		55574..57969	
repeat_region		/note="trigger1 repeat: matches 1..2417 of consensus"	
misc_feature		complement(59285..59816)	
repeat_region		/note="match: GSS: Em:A0377376"	
misc_feature		63327..64290	
repeat_region		/note="LTR5 repeat: matches 1..969 of consensus"	
misc_feature		complement(66506..66624)	
misc_feature		/note="match: GSS: Em:A0390658"	
misc_feature		complement(66507..66624)	
misc_feature		/note="match: GSS: Em:A0718621"	
misc_feature		complement(66617..66697)	
Alignment Scores:			
Pred. No.:	0.00109	Length:	171146
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	9	Gaps:	0
US-09-397-967-16 (1-1099) x AL161450 (1-171146)			
Oy 1003 SeraspValITPserPheG1YValValleutyrgIuleupherhTyr 1018			
Db 104338 TCACATGTTTGGAGCTTGGAGCTGTTCTGTATGATGACTTTCACACTAC 104385			
RESULT 54			
AF173032	AF173032	4744 bp	mRNA
LOCUS	Mus musculus	tyrosine kinase TyrK2 (Tyrk2)	linear
DEFINITION			ROD 15-AUG-1999
ACCESSION	AF173032		mRNA, complete cds.
VERSION	AF173032.1	GI:5733094	
KEYWORDS			
SOURCE			
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;		
JOURNAL	1 (bases 1 to 4744)		
REFERENCE	McCoy,B., Range,A., Lassnig,C., Karaghiosoff,M. and Muller,M.		
AUTHORS	Cloning and characterization of mouse Tyrk2 cDNA		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 4744)		
AUTHORS	McCoy,B., Range,A., Lassnig,C., Karaghiosoff,M. and Muller,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-jun-1999) Animal Production and Genetics, University		
FEATURES	of Veterinary Medicine, Veterinärplatz 1, Vienna 1210, Austria		
source	Location/Qualifiers		
	1..4744		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/chromosome="8"		
	1..4744		
	/gene="Tyrk2"		
	216..3758		
	/gene="Tyrk2"		
	/codon_start=1		
	/product="tyrosine kinase TyrK2"		
	/protein_id="A049423.1"		
	/db_xref="GI:5733095"		
	/translation="MPLGRRALIEDSKADGTEADPLVPTGCLMYLHWPRGEGEP		
	VTFPSLSLAEVCIHIAHKYITPPCINLFAALNAQKWLPPVHILDTSDMNLVY		
	RRERFMHMGNDPEPAVYKRGCGATSSDRAQGLDPSAFEYLFGKGHEF		
	NDVYSIRLDSSEETIHFKNSIGAFILCHILASRGVPLEMARETISFNKCI		
	ROTHQHNVTIRLRILHRYRFRTRIDRFGCHLSQOVVYKYLATLERRLAPR		
	CHLEVDAPERPDCYIONSQCATDPCPELDSGPTHEVYLVGTGSGIOWHPLQ		
	NSGNHGSKSGKKRPKAPAGELTSPDEPWYTCDFODISHVYAKERRVHIIH		

ODNKCLLCLCSQAEALSFVALVDGFRILTDSDSHYLCEHVAAPRLVTSIONGIHPL  
 MDPVOAKIMPEGLYLQWSTSHLRLLTYAHNPXNSNGPRLRKRKPEITGOOP  
 AFVLDGMSFSLGLDRLALOGCSLRAGDDCFPLHXCCLPPRLSNLIMGRSAH  
 TRPLNLSLSFRRHVDITOLSHLGSGCRTRPVGCLLRVCGPDGKDNCGPPEPG  
 TSGQLRLVLYKLDPSHHDLALAFEXASLMSQVSHMLALFHVGCYSGSEITVET  
 VEHGLDWLRKQVGPMTKMWAVQASLALSTLEKNNLHVGVCSRNILLARGL  
 BEGTPFKILSDPGVGOALSHREERERIPWAPRCLSLATLTPEDAPRSPFTIL  
 EICFDEAPLQGRGPEKERYTKKHQLPESSPELATLTROCLTYEDAPMGFGATLL  
 RDLRLQONLVGTSAVNSDSPADPVEHKKYKIKIDLGEGHGVSLVLCYDPTND  
 GTGEVAVKALKEGGCPOLRSGOREITLRLTHEHVKYKCCCEODGEGSVQVME  
 VPLGLSLDYLPRHCVGLAQILLFNOQCEGNAYLHACHYIHRDLAANVLIDNRLV  
 KIDGGLAKAVPEGEHYRVREDGSPVFWAPECLECKRYVASVGVTLVELL  
 TYCDNSQSPHMKFTEILIGHTQGMVLRILTLELGERLPRDRCPCEIYHLMKNCE  
 TEASFRTFONLVPILOTQAEKYCGVPSVFSVC"

BASE COUNT 1036 a 1358 c 1341 g 1006 t 3 others

Alignment Scores:  
 Pred. No.: 0.000584 Length: 4744  
 Score: 15.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.36% Indels: 0  
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AF173032 (1-4744)

QY 333 GluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347  
 DB 1434 GAGGCCCTCTCTTGTGGCCCTGCGCATGGCTATTTCGCTTG 1478

RESULT 55  
 AF052607 7134 bp DNA linear ROD 05-SEP-1998  
 LOCUS Mus musculus non-receptor tyrosine kinase (Tyk2) gene, partial cds.  
 DEFINITION AF052607  
 VERSION AF052607.1 GI:3549252  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 7134)  
 Laasnig,C., Karaghiosoff,M., Range,A., Brem,G. and Muller,M.  
 Direct Submission  
 Submitted (05-MAR-1998) Department of Veterinary Biotechnology and  
 Molecular Genetics, Institute of Animal Breeding and Genetics  
 Veterinary University of Vienna, Veterinaerplatz 1, Vienna 1210,  
 Austria

FEATURES  
 SOURCE  
 Location/Qualifiers  
 1..7134  
 /organism="Mus musculus"  
 /strain="129/Sv"  
 /db\_xref="taxon:10090"  
 /chromosome="8"  
 /issue\_type="liver"  
 <2362..>7072  
 /gene="Tyk2"  
 join(<2362..2479,2782..2929,3200..3363,5578..5944,  
 6629..6829,6915..>7072)  
 /gene="Tyk2"  
 /product="non-receptor tyrosine kinase"  
 join(<2362..2479,2782..2929,3200..3363,5578..5944,  
 6629..6829,6915..>7072)  
 /gene="Tyk2"  
 /note="member of Jak (Janus) family; contains Jak homology  
 regions 7 to 4"  
 /codon\_start=3  
 /product="non-receptor tyrosine kinase"  
 /protein\_id="AAC34580.1"  
 /db\_xref="GI:3549253"  
 /translation="ITPPCLNIFALYNAQKVMVLPNHLDTSDNNLYPRMRFYFN  
 WHGNPQEPAYVRCGFGAETSSDRAGVGQLDLSASFYLFEQGRKHFXNDVYSLRD

LSSFEIHFKNESLGMFLFLHIALSRGVLEPMAREISFKNCIRPSRQIHOH  
 VLTJRLRVRRRRLRARPRHLSQVVMVYLTLETLARFSSERLIPVCHLELAD  
 PERDCTIÖNSGQIADGPELPSGPPTHEVLVGTGIQHPILQTESENGNSGNP  
 HGSRSGRKPKAPKEGHEHLTESPQEPPTHEVDFODLSHVVLKERRVHTHODNCLL  
 CLCSQXELSLVALVDGYFRFLTASSHYLCEHVAAPRLVTSIONGIHPL"

BASE COUNT 1684 a 1824 c 1807 g 1785 t 34 others

Alignment Scores:  
 Pred. No.: 0.00081 Length: 7134  
 Score: 15.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.36% Indels: 0  
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AF052607 (1-7134)

QY 333 GluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347  
 DB 6939 GAGGCCCTCTCTTGTGGCCCTGCGCATGGCTATTTCGCTTG 6983

RESULT 56  
 AF090382 14907 bp DNA linear VRT 05-OCT-2000  
 LOCUS Tetracodon fluvialtilis jak2 tyrosine kinase (JAK2) gene, complete  
 DEFINITION AF090382  
 VERSION AF090382.1 GI:5918519  
 KEYWORDS  
 SOURCE Tetracodon fluvialtilis.  
 ORGANISM Tetracodon fluvialtilis.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorphia; Acanthopterygii; Percormorpha; Tetraodontiformes;  
 Tetraodontidae; Tetraodon.  
 1 (bases 1 to 14907)  
 Leu,J.H., Yan,S.J., Lee,T.F., Chou,C.M., Chen,S.T., Hwang,P.P.,  
 Chou,C.K. and Huang,C.J.  
 Complete genomic organization and promoter analysis of the  
 round-spotted pufferfish JAK1, JAK2, JAK3, and Tyk2 genes  
 DNA Cell Biol. 19 (7), 431-446 (2000)  
 20399330  
 PUBMED 10945233  
 2 (bases 1 to 14907)  
 Leu,J.H. and Huang,C.J.  
 Direct Submission  
 Submitted (03-SEP-1998) Biological Chemistry, Academia Sinica,  
 Taipei P.O. Box 23-106, Taiwan

FEATURES  
 SOURCE  
 Location/Qualifiers  
 1..14907  
 /organism="Tetracodon fluvialtilis"  
 /db\_xref="taxon:47145"  
 2522..13979  
 /gene="JAK2"  
 join(2522..2988,4991..5227,5684..5807,5898..6015,6882..7027,7107..7398,7537..7667,7829..7969,8115..8226,8594..8780,9051..9175,9255..9389,9525..9612,10051..10178,10283..10421,10512..10663,10763..10913,11033..11169,11320..11509,11692..11816,11903..12075,12161..12278,12462..12575,12778..13978)  
 /gene="JAK2"  
 /product="jak2 tyrosine kinase"  
 join(2522..2988,4991..5004)  
 /gene="JAK2"  
 join(5005..5227,5684..5807,5898..6015,6882..7027,7107..7398,7537..7667,7829..7969,8115..8226,8594..8780,9051..9175,9255..9389,9525..9612,10051..10178,10283..10421,10512..10663,10763..10913,11033..11169,11320..11509,11692..11816,11903..12075,12161..12278,12462..12575,12778..12885)  
 /gene="JAK2"  
 /codon\_start=1

/product="ak2 tyrosine kinase"  
 /protein\_id="ra556051.1"  
 /db\_xref="GI:5918520"  
 /translation="MASLAMEETLADACTYHQNGTAKRESEPEARLSYDULHFTNRS  
 /transaltion="MASLAMEETLADACTYHQNGTAKRESEPEARLSYDULHFTNRS  
 GGGATDSSFLSTYPNGPAAWEEIC IDAAKCSISTPYLSFLGIFRRIDDLMEPSPHNQ  
 LDEASADVFEEIETRYEPFGYSGAGSRAAYCAVGLSSPVEDCVMSYLETSQMRNDQ  
 VNGMGRIPSEHIEEIRYEEAGVADLMDTRKAEQLSPDLIDYTMYSKSELPSEIAQIQ  
 COLITVAPADLPFQALHPVAPGVSPPHARPOTELLIMESLEKAFYEPFOVLEISR  
 GGLIIVVADNGIQMOKCEKIDKSDDELTOIDLPDVPYDITIKASKSEAPSEAVTIN  
 KODGKNVLEALSEALSPEVSDIGYRTITTDHNNHLCKEVAAPRLVYATISHGSGPIS  
 LEHAINOLQSGCKNGGIFLQDSKPOENFTFLPEVAYGVERHCOITRSECGQNP  
 LSGRNRSSLHLELCKSKNEVRSVSVPFGSCPPKPAKREASCTLAVCRNKPEV  
 LSPSTNNHNSIQNSPFKIRKEDLEFNESIGGTFPKIRGVAKLEGADVGNKQHEVYM  
 KYDQARNRSESEFPEAKSMQNLSTVHLLILVAYGVCSEENIMQEVKSGESDLYL  
 KKNSSGIVTLPKEILVEYLRPAWPCVPCVNDLNLADKNSGFTGIMELSGSEGRFL  
 LKSDGSIYTLPELVEYLRPAWPCVPCVNDLNLADKNSGFTGIMELSGSEGRFL  
 ASDLNTKKTFLYEDRQLQAPKPKIELANITSCMDYEPFLPRTPAIRITDLSLSPEDD  
 LELTVSDDLIPNRKATGWSGGFSDQPAQREKHLFLDQLGNGNSYKCRIDPP  
 LQDWTGVAVVAKLQHTGATLIDFERELEITLQSHENIVAKYGVGVSAAKRNKIDPP  
 MEYLPESLSDIGTLKNNKREDHKLIVGYSQCKMEYLSQKRYHDLARNLVYS  
 ESRVAKIDFGKTLIDPKDYVYVRGSPSPFWAPSLSEKRSYVADWSGVVYL  
 YELTFTDNNRNSPPTVMSMGNDQGGOLIVHLLIELLKSGSRLPQPLDCTPIHEITME  
 ECMDNDNRPYRSEFKELALSIDLFRHKRF"  
 128686..13978

Alignment Scores:	
Pred. No.:	0.00146
Score:	15.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	1.36%
DB:	5
Length:	14907
Matches:	15
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-397-967-16 (1-1099) x AF090382 (1-14907)

Qy	1003	SeraspValTrrSerPheGlyValValLeuTyrgIuleuPheThr	1017
Db	12204	TCAGATGTCGAGCTTCGGGGTGCTGCTATGAGCTCTTCAAC	12248

RESULT	57
AC027154/c	
LOCUS	172936 bp DNA linear HTG_31-JAN-2001
DEFINITION	Mus musculus chromosome 6 clone RP23-37F10 map 6, WORKING DRAFT
ACCESSION	AC027154
VERSION	AC027154
KEYWORDS	AC027154.3 GI:12621447
SOURCE	HTG; HTGS_PHASE1; HTGS_DRAFT.
	Mus musculus.

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS
1 (bases 1 to 172936)	Birren, B., Linton, L., Nusbaum, C. and Lander, E.	Mus musculus chromosome 6, clone RP23-37F10	Unpublished	2 (bases 1 to 172936)	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.

Boguslavskii, L., Boukhalter, B., Brown, A., Buretkii, G., Camprodoni, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J., S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gaidyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howard, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karacas, A., Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, N., Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGuck, A., McKernan, K., McPheters, R., Melidim, J., Meunier, L., Mhova, T., Miranda, C., Molina, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

O'Neil,D., Ollivar,T.M., Oliver,J., Peterson,K., Pierre,N.,  
 Pisanik,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
 Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talmas,J.,  
 Messiaene,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,  
 Vassiliiev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.,J.,  
 Young,G., Zaitoun,J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (28 MAR 2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jan 31, 2001 this sequence version replaced g11:9368491.

All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 ----- Genome Center -----  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: MIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information -----  
 Center project name: 15743  
 Center clone name: 37\_F10  
 ----- Summary Statistics -----

Sequencing vector: M13; M77815; 49% of reads  
Sequencing vector: pLSD; n/a; 51% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 170054 bases at least Q40  
Consensus quality: 171641 bases at least Q30  
Consensus quality: 172109 bases at least Q20  
Insert size: 198000; agarose-*fp*  
Insert size: 172536; *sum-of-configs*  
Quality coverage: 8.3 in Q20 bases; agarose-*fp*  
Quality coverage: 9.5 in Q20 bases; *sum-of-configs*

```

* NOTE: This is a working draft sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

\* 8836: contig of 8836 bp in length  
 1 8837 gap of 100 bp  
 8837 93243: contig of 84307 bp in length  
 93244 93343: gap of 100 bp  
 93344 110724: contig of 17881 bp in length  
 110725 110834: gap of 100 bp  
 110735 110835: contig of 56244 bp in length  
 110835 167748: gap of 100 bp  
 167749 167848: gap of 100 bp  
 167849 172936: contig of 5088 bp in length

FEATURES	Location/Qualifiers
source	1. .172936

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="6"
/map="6"
/clone="RP23-37P10"
/clone_lib="RPci-23 Female Mouse BAC
1.8836
misc_feature

```

misc\_feature

misc\_feature

misc\_feature

misc\_feature

BASE COUNT 44231 a 43573 c 42293 g 42434 t 405 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.0104 Length: 172936  
Score: 15.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.36% Indels: 0  
DB: 2 Gaps: 0  
US-09-397-967-16 (1-1099) x AC027154 (1-172936)  
QY 333 GUAUAlauserPheValAlaLeuValaAspGlyTyrPheArgLeu 347  
|||||  
Db 128207 GAGGCCCTGCTTGTGGCCCTGTCGATGCTATTTCGCTTG 128163  
RESULT 58  
LOCUS AF091238 8788 bp DNA linear VRT 05-OCT-2000  
DEFINITION Tetradon fluvialilis JAK3 tyrosine kinase (JAK3) gene, complete  
ACCESSION AF091238  
VERSION AF091238.1 GI:5739117  
KEYWORDS Tetradon fluvialilis.  
SOURCE Tetradon fluvialilis.  
ORGANISM Tetradon fluvialilis.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Tetraodontidae; Tetraodon.  
1 (bases 1 to 8788)  
Leu,J.H., Yan,S.J., Lee,T.F., Chou,C.M., Chen,S.T., Hwang,P.P.,  
Chou,C.K. and Huang,C.J.  
Complete genomic organization and promoter analysis of the  
round-spotted pufferfish JAK1, JAK2, JAK3, and TYK2 genes  
DNA Cell Biol. 19 (7), 431-446 (2000)  
20399330  
PUBMED 10945233  
REFERENCE 2 (bases 1 to 8788)  
AUTHORS Yan,S.J. and Huang,C.J.  
TITLE Direct Submision  
JOURNAL Submitted (09-SEP-1998) Institute of Biological Chemistry, Academia  
Sinica, PO Box 23-106, Taipei, Taiwan, Republic of China  
FEATURES  
SOURCE location/Qualifiers  
1..8788  
/organism="Tetradon fluvialilis"  
/db\_xref="taxon:47145"  
2163..8636  
/gene="JAK3"  
mRNA join(2163..2449,2652..2913,2988..3105,3170..3287,3378..3523,3596..3899,3984..4103,4196..4353,4492..4606,4710..4890,5039..5163,5321..5449,5633..5717,5821..5945,6075..6213,6328..6479,6618..6762,6917..7044,7123..7312,7387..7511,7610..7782,7993..8110,8194..8307,8384..8503)  
/gene="JAK3"  
/product="JAK3 tyrosine kinase"  
join(2715..2913,2988..3105,3170..3287,3378..3523,3596..3899,3984..4103,4196..4353,4492..4606,4710..4890,5039..5163,5321..5449,5633..5717,5821..5945,6075..6213,6328..6479,6618..6762,6917..7044,7123..7312,7387..7511,7610..7782,7993..8110,8194..8307,8384..8503)  
/gene="JAK3"  
/codon\_start=1  
/product="JAK3 tyrosine kinase"  
/protein\_id="AAB50352.1"  
/db\_xref="GI:5739118"  
/translation="MNIAREFAFLVIRDRGSGTSSSSAGSLQVHLVFPATKAT  
TTHSGOIAENVCIOAGKKSGLIPVIGLFIASADLSFWPPSHMNTENILAVC  
FRVFFGNMFDGPRACRYSLTRDTSPVLDHSLIAYFNLRDPTFSSGSLSP  
LTPDQICGLAVDLIRAKESQSSVKDLCKSVSYSCLPKSHIRVIERNWERRI  
RNLKLFKRFSGCAVDNCLKMLVLTALAGIOPISGSEFLHPSPSSILNSAVSL  
VQVTEIGTISGEGVWQFCEKIDTISIKRICREOVFNDSRMVTITRKDAC

LEVAFOSLKEALSLVSDYGFRLTDTSTHYFCODTAPPSILGINKHCHPINSERA  
VNKLKRWSEDEGTFLLRKSPDYGEFFLTVAAGTPLGLDKDLVINKKHSLPSS  
KFSFLKELFYVYOONKLLADVAEPVLSRCCPPSPSOELTNLIIRNSFPVLAOLSP  
EGNKLHIHFOPIMYKEDLOMGESHGOSFTFLPFKYSKNSISEGKRVPLKEDLVH  
RNLMSFEFASLSMSPSHKHLITLVCTAVAKNIMQEPREYALDYLKRGSVSV  
SMKLDVAKOLASVITFLEONITVGNICAKNLLARSDPSQIIPPIKISDPISVL  
MLGADVDRITPWAPEVLASPENULTLESDKWSGATLWELFNNGNPNLGLDLMKF  
YESQQLPESQWTELDLINOCDYQADFRSCRSIIRQLSITSDYVILFATEPDT  
QPPAMRALTPSOQDOTVEERHLRYISILGNGSVLELCCYDPLGNGTELVAVKL  
QPSKSSMEDFKKEVTLISVHCYISIKYGVCSMGRISMGVTEYLPYSGILGYE  
NNRHKVDTRMLLEFASOICGMEYLOSIRAFVHRDLARNILVASESLVKTADGTRIT  
IPCKEYRYVTPGESPIFWAPPSINSESRSHSDVSGVVALTELSYCDINSNK  
RLYKQOIGHVQTPSTISLHNLILKSNMRLPAPPDCAKVISLMECEMLDFMERPCF  
SFLNQLEIVIMODERDSFKG  
BASE COUNT 2291 a 2068 c 2109 g 2320 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.00903 Length: 8788  
Score: 14.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.27% Indels: 0  
DB: 5 Gaps: 0  
US-09-397-967-16 (1-1099) x AF091238 (1-8788)  
QY 1003 SerAPVAlTPSPserPheGlyValValLeuTyrGluLeuPhe 1016  
|||||  
Db 8036 TCCGACGTTTGAGATTGTGAGATGTTTACGAGCTCTTC 8077  
RESULT 59  
LOCUS AR002620 41 bp DNA linear PAT 04-DEC-1998  
DEFINITION Sequence 9 from patent US 5741899.  
ACCESSION AR002620  
VERSION AR002620.1 GI:3964174  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 41)  
AUTHORS Capon,D.J., Tjian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.  
TITLE ChimERIC receptors comprising Janus kinase for regulating cellular  
pro liferation  
JOURNAL Patent: US 5741899-A 9 21-APR-1998;  
FEATURES location/Qualifiers  
SOURCE 1..41  
/organism="unknown"  
BASE COUNT 7 a 9 c 15 g 10 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.00116 Length: 41  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 6 Gaps: 0  
US-09-397-967-16 (1-1099) x AR002620 (1-41)  
QY 1 MetAlaProBserGluGluThrProLeuIleProGln 13  
|||||  
Db 40 ATGGCACCTCCAGTGAGAGACACCTCGATCCCTCAG 2  
RESULT 60  
LOCUS AR099670 41 bp DNA linear PAT 14-FEB-2001  
DEFINITION Sequence 9 from patent US 6077947.  
ACCESSION AR099670  
VERSION AR099670.1 GI:12809436  
KEYWORDS

SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 41)  
AUTHORS Capon,D.J., Tian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.  
TITLE DNA encoding an intracellular chimeric receptor comprising Janus kinase  
JOURNAL Patent: US 6077947-A 9 20-JUN-2000;  
FEATURES Location/Qualifiers  
source 1..41  
BASE COUNT 7 a 9 c 15 g 10 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.00116 Length: 41  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AR099670 (1-41)

OY 1 MetalapProserglutThProleuileProgin 13  
Db 40 ATGGCACCCTCCAGTGAAGACACCTCTGATCCCTCAG 2

RESULT 61  
AR002619 45 bp DNA linear PAT 04-DEC-1998  
LOCUS AR002619  
DEFINITION Sequence 8 from patent US 5741899.  
ACCESSION AR002619  
VERSION AR002619.1 GI:3964173  
KEYWORDS  
ORGANISM Unknown.  
SOURCE Unknown.  
REFERENCE 1 (bases 1 to 45)  
AUTHORS Capon,D.J., Tian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.  
TITLE Chimeric receptors comprising janus kinase for regulating cellular proliferation  
JOURNAL Patent: US 5741899-A 8 21-APR-1998;  
FEATURES Location/Qualifiers  
source 1..45  
BASE COUNT 11 a 16 c 10 g 8 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.00125 Length: 45  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AR002619 (1-45)

OY 1 MetalapProserglutThProleuileProgin 13  
Db 6 ATGGCACCCTCCAGTGAAGACACCTCTGATCCCTCAG 44

RESULT 62  
AR099669 45 bp DNA linear PAT 14-FEB-2001  
LOCUS AR099669  
DEFINITION Sequence 8 from patent US 6077947.  
ACCESSION AR099669  
VERSION AR099669.1 GI:12809435  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 45)  
AUTHORS Capon,D.J., Tian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.  
TITLE DNA encoding an intracellular chimeric receptor comprising Janus kinase  
JOURNAL Patent: US 6077947-A 9 20-JUN-2000;  
FEATURES Location/Qualifiers  
source 1..41  
BASE COUNT 7 a 9 c 15 g 10 t  
ORIGIN

REFERENCE 1 (bases 1 to 45)  
AUTHORS Capon,D.J., Tian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.  
TITLE DNA encoding an intracellular chimeric receptor comprising Janus kinase  
JOURNAL Patent: US 6077947-A 8 20-JUN-2000;  
FEATURES Location/Qualifiers  
source 1..45  
BASE COUNT 11 a 16 c 10 g 8 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.00125 Length: 45  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AR099669 (1-45)

OY 1 MetalapProserglutThProleuileProgin 13  
Db 6 ATGGCACCCTCCAGTGAAGACACCTCTGATCCCTCAG 44

RESULT 63  
AY078944 1592 bp mRNA linear PLN 09-MAR-2002  
LOCUS AY078944  
DEFINITION Arabidopsis thaliana At2g28940/F8N16.23 mRNA, complete sequence.  
ACCESSION AY078944  
VERSION AY078944.1 GI:19310416  
KEYWORDS  
SOURCE FLI\_CDNA.  
ORGANISM Arabidopsis thaliana.  
SOURCE Arabidopsis thaliana.  
REFERENCE 1 (bases 1 to 1592)  
AUTHORS Shin,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bower,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
TITLE Arabidopsis cDNA clones  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1592)  
AUTHORS Shin,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bower,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
TITLE Direct Submission  
JOURNAL Submitted (08-FEB-2002) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
COMMENT  
There is no suitable ORF in this cDNA clone.  
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFI cDNAs (RAFI cDNA: 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.  
The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFI cDNAs: Shin,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bower,L., Chang,E., Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B.,



Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Shim, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

## FEATURES

Location/Qualifiers  
1. 1592  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="RAF109-69-D18(R21702)"  
/note="ecotype: Columbia"

BASE COUNT 462 a 270 c 403 g 457 t

## ORIGIN

## Alignment Scores:

Pred. No.: 0.0217 Length: 1592  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 8 Gaps: 0

US-09-397-967-16 (1-1099) x AY078944 (1-1592)

Oy 1003 SeraspValTrrpSerPheGlyValIleuTyGluLeu 1015  
|||||  
Db 948 AGTGATGTATGAGCTTTGGGAGTGTCTATGAACTG 986

## RESULT 64

TFU53213 16977 bp DNA linear VRT 03-FEB-1998  
LOCUS Tetradon fluvialilis JAK1 tyrosine kinase (JAK1) gene, complete  
DEFINITION  
cds.  
ACCESSION U53213  
VERSION U53213.1 GI:2828795

## SOURCE

ORGANISM Tetradon fluvialilis.  
Tetradon fluvialilis.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetradon.  
1 (bases 1 to 16977)  
Leu, J.H., Chang, M.S., Yao, C.W., Chou, C.K., Chen, S.T. and Huang, C.J.  
Genomic organization and characterization of the promoter region of  
the round-spotted pufferfish (Tetradon fluvialilis) JAK1 kinase  
gene

## REFERENCE

AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
9434151  
2 (bases 1 to 16977)  
Leu, J.H., Chang, M.S., Chang, G.D. and Huang, C.J.  
Direct Submision  
Submitted (01-APR-1996) J.H. Leu, Inst. Biological Chemistry,  
Academia Sinica, P.O. Box 23-106, Taipei, Taiwan 106, Republic of  
China

## FEATURES

Location/Qualifiers  
1. 16977  
/organism="Tetradon fluvialilis"  
/db\_xref="taxon:47145"  
/clone="pf11"  
5236. 16216  
/gene="JAK1"  
join(5236..5440,6078..6201,7605..7758,7987..8138,  
8245..8513,8714..8823,8930..9115,9233..9390,9592..9706,  
9836..10034,11673..11782,11920..12072,12201..12288,  
12367..12494,12615..12753,13464..13615,13691..13844,  
14295..14383,14525..14717,14786..14910,15297..15469,  
15563..15680,15876..15986,16109..>16216)

## CDS

/gene="JAK1"  
join(5236..5440,6078..6201,7605..7758,7987..8138,  
8245..8513,8714..8823,8930..9115,9233..9390,9592..9706,  
9836..10034,11673..11782,11920..12072,12201..12288,  
12367..12494,12615..12753,13464..13615,13691..13844,  
14295..14383,14525..14717,14786..14910,15297..15469,  
15563..15680,15876..15986,16109..16216)

/note="protein tyrosine kinase"  
/codon\_start=1  
/product="JAK1 tyrosine kinase"  
/protein\_id="AAC00207.1"  
/db\_xref="GI:2828795"

/translation="MSTQVWELSNOLGCKRRGRKAHSSPPSPFGEIHTPT  
VHQLTSGCTTAELEAARKCSISPLCNLFALDEKNNMVPNNYETVEELS  
LKLHYRMFEYFRNMHNGTEGSSVWRHICSLKRGSLSPQKIPVGTPLDASLELY  
QGSHSFHKLGLVPRKSHSDTEQHEIENECLEMAVALATHTNVCAGPSHDISYKRF  
ETLNRSIKORSFETPIRINNVKFKFSEFHQRTVSDNSITPYDLKIKYATLEGTSG  
LGSEVFEPISLVAOEGELGNGVYRYFNQGOAVKTSRQVAVTGTGISMK  
KPEMALTSKDKSKSKSKADCKOONRKRKRETCMEYFCDFYETITAVIADKTVIT  
VRODNMGELOMASRAESLRYALVDYFRITVDHHYLCQDVA PASVVRNISCCH  
PISTIEAHKLKRGNEGYILRMSCDFEYIVITVYCELELDKESRPAQYKNFOI  
EESDGEFLYGTDCRPTLMLEHLSQSRLTDNLOFOLRVCPQPPREISNLVYT  
KDRVLTSPSPDQSLSPHRLKEDIQEEHIGRTGRTNIYSGLTVKSEEDAGYS  
SFOEVKVLKLSYGHNDISLAPEFETSMRQVSKHKLIVLLGYCAVHONIMVEFY  
QLGPLEVEMRROGSLSPMKFOYAKOLASLSYLEDKILYHGVCCKNILLADGCG  
TDEGGPFIKLSDPGIPITVLTRECVRIPIAECYKVNYSALSVADKMGFTTME  
ICVGEVPLKPKKITEQKERYETECOLATPDCTELALMTHCNVADPKRPFERATV  
RDIIMLEQNPSTIKPKPAEVDVPEFIDIKALFHEHIVYKIGICEGGQARLPDQK  
ELVAVKSLKPEQDQSTNLSREIDIKALFHEHIVYKIGICEGGQARLPDQK  
MGSLLKDYLPRIHKKDTSLATILSYVQICGMEYIGSNYIHRDLAARNVAVENRMV  
IGDGLTKRSIKDNEGTYTVKDNDSPEVMAPELCKCKFYRLASLDVNSPVTYELIT  
YDSSKSPMTCFLDMIGMTQOMVNLVNLNDGRLLPRPDGCPREYIELMKRCEQ  
KPDRIETFIKIKELSMQOEPQ"

BASE COUNT 4600 a 3563 c 3734 g 3080 t

## ORIGIN

## Alignment Scores:

Pred. No.: 0.144 Length: 16977  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x TFU53213 (1-16977)

Oy 335 LeuserPheValAlaIleuValAspGlyTyrPheargLeu 347  
|||||  
Db 9263 CTGCTCTTGAGCTCTTGATGGGTATTCGCGCTG 9301

## RESULT 65

AC005727 107097 bp DNA linear PLN 11-MAR-2002  
LOCUS Arabidopsis thaliana chromosome 2 clone F8N16 map m154, complete  
DEFINITION  
sequence.  
ACCESSION AC005727  
VERSION AC005727.3 GI:20197435

## SOURCE

ORGANISM Arabidopsis thaliana.  
Arabidopsis thaliana.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1 (bases 1 to 107097)  
Shen, M., Rongning, C.M., Kaul, S., Shao, T.P., Fujii, C.Y., Mason, T.M.,  
Rounsley, S.D., Lin, X., Fraser, C.M., Somerville, C.R. and Venter, J.C.  
Unpublished  
2 (bases 1 to 107097)

JOURNAL  
AUTHORS  
TITLE  
REFERENCE  
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712  
Medical Center Dr., Rockville, MD 20850, USA

REFERENCE 3 (bases 1 to 107097)  
AUTHORS Town, C.D. and Kaul, S.  
TITLE Direct Submission  
JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org  
COMMENT On Apr 18, 2002 this sequence version replaced gi:5598485.  
FEATURES  
Location/Qualifiers  
Source 1..107097  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/map="m154"  
/clone="F8N16"  
misc\_feature 1..3005  
/note="overlap with BAC clone T1P11 (:14595..17599)."  
gene 187..258  
/gene="At2g28730"  
tRNA 187..258  
/gene="At2g28730"  
/product="tRNA-Cys"  
/complement(678..1179)  
/gene="At2g28740"  
/note="F8N16.2; identical to GB:M17133 and  
GB:M17132; supported by full length cDNA: Ceres:32642"  
gene complement(<678..>1179).  
CDS /gene="At2g28740"  
complement(827..1138)  
/gene="At2g28740"  
/codon\_start=1  
/product="histone H4"  
/protein\_id="AAC79580.1"  
/db\_xref="GI:3927823"  
mRNA /translation="MSGRGKGGKLGKGAKRHKVLRDNIQTTPAIRLRARGGV  
KRISGLIYETRGVLAIFLENVITDAYTYTEHARKKVTAMDVYVALKRGGRLLYGGG  
G"  
repeat\_region complement(1191..1211)  
repeat\_region /rpt\_family="AT\_rich"  
1543..1590  
/rpt\_family="AT\_rich"  
gene 2811..3958  
/gene="At2g28750"  
/note="F8N16.3; pseudogene, reverse transcriptase"  
/pseudo  
complement(5356..5643)  
/gene="At2g28755"  
/note="F8N16.4"  
mRNA complement(join(<5356..5421,5539..>5643))  
/gene="At2g28755"  
complement(join(5356..5421,5539..5643))  
CDS /gene="At2g28755"  
/codon\_start=1  
/product="putative dUDP-glucose 4-6-dehydratase"  
/protein\_id="AA15077.1"  
/db\_xref="GI:20197439"  
/translation="WEGIKRLMAGDKSPINIGNGEFSIYELAEYKALIKPVEIK  
IVENIPDDPQRS"  
repeat\_region 6797..6822  
/rpt\_family="AT\_rich"  
gene complement(7653..9826)  
/gene="At2g28760"  
/note="F8N16.5"  
complement(join(<7653..7835,7913..8017,8271..8337,  
8478..8571,8665..8742,8834..8987,9083..9174,9261..9435,  
9743..>9826))  
/gene="At2g28760"  
complement(join(7653..7835,7913..8017,8271..8337,  
8478..8571,8665..8742,8834..8987,9083..9174,9261..9435,  
9743..9826))  
CDS /gene="At2g28760"  
/codon\_start=1  
/product="putative nucleotide-sugar dehydratase"  
/protein\_id="AAC79582.1"  
repeat\_region /db\_xref="GI:3927825"  
/translation="MASNSNGCTTTTKPPMPSPLRNSKRFQSNRIILVTGAGFIS  
HLYDKIMOEKNEVIVADNPYFGSKDNKKWIGHRFELIRIDVTEPLEFVDOYHL  
ACPAAPDIFKYNPVYTIKKNVIGTINMGLAKRGVARIILITSTSVGDPLVHPTES  
IWNVNPVIVRSCTDEGKRVATLDFPDHRRQIGIRLARIENYGPMMNIDGGVVS  
NFIAGLRREALTVOKPGTQTSFCYSVDWVGLMWDGDTGFINIGNGEFLNVE  
IAETVYELIKPVEIKWENTPDDRRKRPDISKAEVAGWEPKVKLEGLPLMEDF  
RLRIQVPRK"  
repeat\_region 10264..10324  
/rpt\_family="(GAA)n"  
gene 10935..11006  
/gene="At2g28770"  
tRNA 10935..11006  
/gene="At2g28770"  
/product="tRNA-His"  
gene complement(11283..14459)  
/gene="At2g28780"  
/note="F8N16.7; predicted by genscan and genefinder"  
complement(join(<11283..12968,13077..13244,13929..>14459))  
mRNA /gene="At2g28780"  
complement(join(11283..12968,13077..13244,13929..14459))  
CDS /gene="At2g28780"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AAC79583.1"  
/db\_xref="GI:3927826"  
/translation="MTERGRAMWRTCLASAFRTALACTVGSATLYGPEWIRNVAFP  
AFSYVVIILITPDATLGTDLRGCMALVATGCSVGAIVTLKLRPARLTAETTLAA  
ALAAFYVILPNSSTHLVAKRIALGOVLIVIGYIKGATDVMHQLQVASTALGV  
ACVALLVPLPRTACGEVSKCEIGQNTTTRVRIKMAFCGSDSDMSAVASQARVL  
ARSSSLYOTIKRQESVSMTEPERPKITMNOVNDGKLGSMETALGEMVYASK  
SPILPSSULGAEVKEDIKNIOERTYISIKVNNNSQSPVPESDPKRPDCLTLEIP  
GTPQDLPYFELFCIRLETTIILAKPEKNVYLEKFTKTRYSISMDSKIMPALK  
SLSLGAILLGSMESKPNGWAGLPVAVFAARXETFTVTKAKGATIGVYVGMG  
CEVQKFLVYRFLSLPLPFLSSFLSRKMYGQAGISALGAVILLGKNGPSEF  
AIEPRIETFGILCSIMVELVQPTRAANIAKIEISRSFHAIECSLFGAASKADI  
MESOKLRSHLNEIKKETAHAHPEFWSPEFSCYEKLFSLSKMDLILPFGYAI  
GFLGEGQKTRSPCKKEILSNVDDKLSIESITGLKSFEETILKSLDALKEALAKS  
DNTSWDIEKLTNPSPFSFVSPSEPKIITFTYLOHCSVADGLFRVEDEEPEVYKS  
EVALSLCAIGFCYERIGKETRELEENKVEVQSENSHVNLHEISCKIRSLXK"  
repeat\_region complement(13246..13298)  
/rpt\_family="AT\_rich"  
15206..15240  
/rpt\_family="AT\_rich"  
repeat\_region complement(16081..16226)  
/rpt\_family="(TATA)n"  
repeat\_region complement(17388..17489)  
/rpt\_family="AT\_rich"  
repeat\_region complement(17827..17848)  
/rpt\_family="AT\_rich"  
repeat\_region complement(18719..18749)  
/rpt\_family="AT\_rich"  
repeat\_region 21595..21648  
/rpt\_family="AT\_rich"  
repeat\_region 22793..22907  
/rpt\_family="AT\_rich"  
repeat\_region complement(23101..23146)  
/rpt\_family="AT\_rich"  
repeat\_region complement(23479..23502)  
/rpt\_family="AT\_rich"  
repeat\_region complement(23679..23701)  
/rpt\_family="AT\_rich"  
gene complement(25689..26616)  
/gene="At2g28790"  
/note="F8N16.8; supported by full length cDNA: Ceres:2058"  
complement(<25689..>26616)  
/gene="At2g28790"  
complement(25848..26597)  
/gene="At2g28790"  
/codon\_start=1  
/product="putative thaumatin"  
/protein\_id="AAC79584.1"  
/db\_xref="GI:3927827"

/translation="MAKTSILPLAASPELLISFSSAVDTSLFTLVNCPETVPALIO  
PNAGHVLKGGKGFALPTFTSRFNVPTTHSSGRIMRTWCABHNGKFSCLTGGCGRIL  
ECNGLGAPASLAOFDLHGHGHDHDSVGVSLVDGYNPMPTVTPREHGVCYVYVCR  
EDLITKCPALLOYRSHSGHVNACKSCCEAFHDELCCRGHNSPNTCKSHSLFFKX  
ACPSSTFAHDSLSLMDHDCASPRELVIRICH"  
complement(27729..30415)  
/gene="At2g28800"  
/note="F8N16.9: identical to GB:U89272: supported by full  
length cDNA: Ceres:31457"  
mRNA  
complement(join<27729..28078,28179..28328,28637..28740,  
28833..28940,29082..29186,29288..29361,29446..29506,  
29584..29658,29769..29879,29968..>30415))  
/gene="At2g28800"  
complement(join(27853..28078,28179..28328,28637..28740,  
28833..28940,29082..29186,29288..29361,29446..29506,  
29584..29658,29769..29879,29968..30342))  
/gene="At2g28800"  
/codon\_start=1  
/product="chloroplast membrane protein (ALBINO3)"  
/protein\_id="AAC79585.1"  
/db\_xref="GI:3927828"  
/translation="MARVLVSSPSFSPFSGPLIKPSSSLRHSGVGSGGTAOPLPYRSMN  
NKLFSTVRSFSLIPPEHGLDSDVIGAJFTRAESLITLTLADAAYGASVYVTD  
SSAVOKSGMFGFISDAMELVILKIDGLSAYHVPAYGFAITLITLYKATTPYTK  
OOVESTLAMONLOPKTKAQORAGNDEIQLTSLIKQAGVNPAGLCLPLATIPV  
WGLVYALSNVANEGLFTEGFMTIPSLGSPITSAARQSGSISWLFEPVGHPLPMY  
DTVAIVLVLPLIASQVSMETMRPQTPDPAOKMLVFKFLPLMIGFALSVPSGL  
SIYWLNNVLSIAQOYLRKLGAKRNMENASKITISAGRAKRSIAQPDADGERFOL  
KDEKRSKKNKAVAKDTVELVEESQSESGSDDEEEERAGALASSTSKPLPEVGO  
RRSKRSRRRTY"  
complement(30388..30418)  
repeat\_region  
/repeat\_id="AT\_rich"  
/rpt\_family="AT\_rich"  
/rpt\_family="AT\_rich"  
Alignment Scores:  
Pred. No.: 0.629 Length: 107097  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 8 Gaps: 0  
US-09-397-967-16 (1-1099) x AC005727 (1-107097)  
QY 1003 SetAspValTrpSerPheGlyValValLeuTyrGluLeu 1015  
|||||  
Db 98537 AGTGATGTATGGAGCTTGGGCTAGTGTATGAACTG 98499  
|||||  
RESULT 66  
AC005315/c 108393 bp DNA linear PLN 11-MAR-2002\*  
LOCUS  
DEFINITION Arabidopsis thaliana chromosome 2 clone T914 map m154, complete  
sequence.  
ACCESSION AC005315  
VERSION AC005315  
KEYWORDS GI:20197356  
SOURCE HTG  
ORGANISM Arabidopsis thaliana.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 108393)  
Rounsley,S.D., Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Mason,T.M.,  
Shen,M., Ronning,C.M., Fraser,C.M., Somerville,C.R. and Venter,J.C.  
JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712  
TITLE Medical Center Dr., Rockville, MD 20850, USA  
JOURNAL 2 (bases 1 to 108393)  
REFERENCE Direct Submission  
AUTHORS Lin,X.  
JOURNAL 2 (bases 1 to 108393)  
REFERENCE Submitted (09-MAR-2000) The Institute for Genomic Research, 9712  
AUTHORS Medical Center Dr., Rockville, MD 20850, USA  
REFERENCE 3 (bases 1 to 108393)  
AUTHORS Town,C.D. and Kaul,S.

TITLE Direct Submission  
JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712  
COMMENT Medical Center Dr., Rockville, MD 20850, USA, cdromelgr.org  
FEATURES On Apr 18, 2002 this sequence version replaced gi:598462.  
SOURCE Location/Qualifiers  
1..108393  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/map="m154"  
/clone="T914"  
1..15649  
/note="overlap with BAC clone F8N16  
AC005727:91449..107097.".  
1108..1131  
/rpt\_family="TAAA)n"  
complement(3674..3724)  
/rpt\_family="AT\_rich"  
4287..6301  
/gene="At2g28930"  
/note="T914.1: contains a protein kinase domain profile  
(PDOC00100).  
join<4287..4329,4769..5088,5179..5314,5404..5546,  
5623..5746,5829..>6301)  
/gene="At2g28930"  
join(4287..4329,4769..5088,5179..5314,5404..5546,  
5623..5746,5829..6301)  
/gene="At2g28930"  
/codon\_start=1  
/product="putative protein kinase"  
/protein\_id="AAC33221.1"  
/db\_xref="GI:3461835"  
/translation="MGCLISQAQIKRVSFGASPKYSSSANDSLSSKSSSVIRNPRT  
EGEILQSPNLKSFTRAEILKAATRNRPDSVLGEGFGSVFEGWIDEQTLTKSKPGTG  
VTAARKLQDQMQOEWLAEVNYLGQFSHPNLVXKLIGYCLEDEHRLTYEFEPMSGL  
ENHLEFRSGSYOPSLWTLRLVVALGAAGLAFLNHAETSVYRDKTSNIIJDSEYNA  
KLSPGLAKDPTGDKSHVSTRIMGTGYAAPEYLATGHLTTKSDVYSYGVALLLEVIS  
GRRAVDKMRPPEGOKLVEMARPLANKRKLRFVIDNRLODQYSMECKKVAITLRLC  
TEFIRKLRNMNEVYSHLEIQTLINEAGGRNIDMQRRRRRSDSVAILQKNAGFARO  
TAVCYIATAYRPSDSPLEFY"  
complement(6589..8414)  
/gene="T914.2: contains a protein kinase domain profile  
(PDOC00100).  
complement(join<6589..7152,7222..7344,7425..7688,  
7977..>8414))  
/gene="At2g28940"  
complement(join(6589..7152,7222..7344,7425..7688,  
7977..8414))  
/gene="At2g28940"  
/codon\_start=1  
/product="putative protein kinase"  
/protein\_id="AAC33222.1"  
/db\_xref="GI:3461836"  
/translation="MCKFHFTNGDKRRTTIIIEGSGGDSVYSRASHLSMARSLVAS  
STSDPTPRSEFSDMSFSEPRLTFFPKLSORWIGLVLPENDLKVFFPKELKATKGF  
NRGLIGEGGCGYRGVVDSDSNGDSKTI NVAVKOLNROGLOGHKEVNEVFLCY  
VHNPIYIKVINGYCADDDERGMRLYIELMCKSLDEHLVGRVSVSLPMMRKLKIAQ  
DAAGGLAVIEHMDFOITRFDKSSNILLDERFAKLSDFGLAQGPBEIGVHVSIV  
VGTVGAYAPREVQGLTKARSDWSFVLYELITGRADVDRNRPDEOKLLEVKPY  
VSDSKRPHLLVDPRLGQYCKMSKVCVVAALANKCKOPKSRPKMEVYSLGRITD  
EEANRPPVPAVDETEIITKALNESEPELKQSGSSRRKVLVDLRDMNLKSLSKID  
WRNWTPELVRTW"  
8721..8765  
/rpt\_family="AT\_rich"  
10356..10398  
/rpt\_family="AT\_rich"  
complement(11426..13331)  
/gene="At2g28950"  
/note="T914.3: identical to GB U30480: supported by full  
length cDNA: Ceres: 17914"  
complement(join<11426..11885,12560..12878,13074..>13331))  
mRNA

```
repeat_region
CDS
/gene="At2g28950"
/complement(11460..11490)
/rpl_family="AT_rich"
/complement(join(11576..11885,12560..12878,13074..13218))
/gene="At2g28950"
/codon_start=1
/product="expansin Atx6"
/protein_id="AAC33223.1"
/db_xref="GI:3461837"
/translocation="MAMIGLVLVLTITLALSEARIPGVYNGGCMETAHATFYGGSDA
SGTMGACGYNTLSGCGVNTALSTALFNNGSCGACCELCASDPKCHGSPSI
FITATNCPNPAFSDNGWCNPRPHDLAMPFLIAEYRAIGVIFVSRPAPCR
RGITRTPNGRPNLYLVNTNACAGNIVRLGVGTGTHSMWTMSRMNGQMOSNVLY
GQSLSFVYSSDRSSTSMNAPAMKFCOTFMCKNRV"
repeat_region
16013..16090
/rpl_family="AT_rich"
17619..17647
/rpl_family="(CAAC)n"
/complement(17650..22083)
/gene="At2g28960"
/nc="T914.4"
/complement(join(<17650..18202,18290..18479,18556..18618,
18720..18846,18923..19155,19260..19417,19497..19574,
19678..19749,19937..20008,20105..20249,20591..21063,
21158..21686,21993..22083))
/gene="At2g28960"
/complement(join(17794..18202,18290..18479,18556..18618,
18720..18846,18923..19155,19260..19417,19497..19574,
19678..19749,19937..20008,20105..20249,20591..21063,
21158..21686,21993..22083))
/gene="At2g28960"
/codon_start=1
/product="putative receptor-like protein kinase"
/protein_id="AAC33224.1"
/db_xref="GI:3461838"
/translocation="WEGKROKLVLFIFGALITLHVQAPDQKGFISLDCGLVNES
PYDPRGLTSSADFTLSLRGEGADNITYIKQYDLFPFGINCLNLYEVS
INYLRAFQGYNGDLVYKFDLWPNMIVADLEFGREITLYITLQICL
VKGTSTIPMISTLELRPLRNDSYLTFGLIYRAVSSSTGIRPDDIPRKMD
RYNEFEDVNTLVNRSSSPQVPEAVSRMGITPENALPRFYSLDDSDGKLVYR
HFAEQALRGNETRDEDEDIQSAVSPMOSDAPYFNISPKKSGSGCYLKLVR
TPRSTLPLPISIAEAKYVDPRPYAETNNDVAAKMDIAPGKIMISMGQPCVPEL
KWEIDKCYTNSKSTPRITISLDSRGKGYAIPFOUTELRKIDNSNNSFTGVE
FLASKRSSTINLNMNDLTGLPKLLDREKNGKLTIOGNPKLNDNSCKNNQOTY
IYPVAVASVYLIIIAVLILILVEKKRRPTQVDS.LPTVHGIPNPSIFGTQKRTYS
EVEALTDNFERVILGGGVYHGIINTOPAVLILSOSSVQKKEKKAVERLLYR
HHVNLSLVGCODESNLALYEAPNGDKOHLSEKSGSLKMSRLKIVETKQGL
EYHTGCKPWHVHRDYKTNLILDEHPQAKLADGLSRSPVGETHSTAVAGPGY
LDPEYTRNRLNEKSDYSPGIVLEITTSRPIVQOTREKPIAMVGYMLTKGDIEN
VVDPRLDYEFTSVKALDEIAMSVCNPSSEKRPWISOVTNELKCOLLENSKRYRE
DMGSRSSVEMSTSTFTEINPKAR"
repeat_region
22868..22909
/rpl_family="AT_rich"
22951..22971
/rpl_family="AT_rich"
23635..23655
/rpl_family="AT_rich"
23635..23655
/gene="At2g28970"
/nc="T914.5"
/join(<23655..23724,23844..24035,24325..24379,24788..25263,
25348..25492,25589..25660,25758..25829,25929..26006,
26187..26344,26421..26671,26769..26895,26994..27059,
27203..27382,27491..>27899)
/gene="At2g28970"
/join(23655..23724,23844..24035,24325..24379,24788..25263,
25348..25492,25589..25660,25758..25829,25929..26006,
26187..26344,26421..26671,26769..26895,26994..27059,
27203..27382,27491..27899)
/gene="At2g28970"
/codon_start=1
/product="putative receptor-like protein kinase"
/protein_id="AAC33225.1"
/db_xref="GI:3461839"
/translocation="MMSHLMLAIGTFAVIVAGAKOKEGFSIDCGPIESPYDPST
```

```
GLTFSTDSFTIOTGESGRVDELNKIFRRPYLTLRYFPEGRKNCISLRNSFRVHCSTSD
SEIRYDDSDRWVWYPPFSSSEYITTSININNSDPEEIPKAKSAATPKNASAPLI
ITKPPRNSAEVYFYIAPRAIOTLANERREPIVKEPNTSASFPKLELLEFTS
GPVQOCSODGNCILVTRPNSTLPLNALAEATIIIEFPOLETSLSYVNAIKTKKAYR
LSKTSWQSDCLPDELSEWENLRCSYSTNSPPKIIISLNASASITGSLDSVQNLTOI
OEIDLNNNSLTGLVPSFLANIKSLSLDLSGNNETSVQTLDRKEGLVLRLEGN
ELCKFSKNSPKKRRKGLLVPIYASISSVLIVYVVALEFVLRKKMPSDAQAPSLVE
DVGAKHNSSESPVSKKIRFAYFEVQETNNFORVLGEGGFQVYHGCNAGTOOAVKL
LSOSSSGQYKHFKAEEILMRVHKNLVSLVGCSDGDLALIEYMPNGDLKOHLSG
KRGCFVTSMSRIRAVADALGLEYLHTGCKPWHVHRDYKTNLILDEHPQAKLADFG
LSSSFPLENTHVSTVYAGTGLDEPYIOTNNLTKSDYSPGIVLEITTRPITIO
OSREKPHLEWVGFIVRTGIDIGNVDPNLHGAYDVGSVMKATIELANSCVNISSARPS
MSQVSDLKXCVISENSRTESEPMSSMSIEFSWGLDVEVIPKAR"
repeat_region
/rpl_family="AT_rich"
28787..28843
/rpl_family="AT_rich"
28870..28926
/rpl_family="AT_rich"
/complement(28945..28964)
/rpl_family="(A)n"
/complement(29072..34092)
/gene="At2g28980"
/nc="T914.6"
/complement(join(<29072..32065,32249..32562,32811..34092))
/gene="At2g28980"
/complement(join(29072..32065,32249..32562,32811..34092))
/gene="At2g28980"
/codon_start=1
/product="putative non-LTR retroelement reverse
transcriptase"
/protein_id="AAC33226.1"
Alignment Scores:
Pred. No.: 0.635 Length: 108393
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.18% Indels: 0
Gaps: 0
US-09-397-967-16 (1-1099) x AC005315 (1-108393)
QY 1003 Seraspvaltrpserphoglyvalvalleutyrgulneu 1015
|||||
Db 7089 AGTGATGATGAGAGCTTGAGGTGATGTAAGTGAAGTAC 7051
|||||
RESULT 67
LOCUS I44512 151 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 11 from patent US 5635177.
ACCESSION I44512.1 GI:2469225
VERSION I44512.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 151)
AUTHORS Bennett,B.D., Goeddel,D. and Matthews,W.
TITLE Protein tyrosine kinase agonist antibodies
JOURNAL Patent: US 5635177-A 11 03-JUN-1997;
FEATURES
source
1..151
/organism="unknown"
BASE COUNT 37 a 44 c 41 g 29 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0311 Length: 151
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
```

DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x I44512 (1-151)

OY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952  
 |||||||||||||||||||||||||||||||||||

Db 1 GTGCACAGAGATCTCGCGCTCGAACAATCCTGCTC 36

RESULT 68  
 AB012713 187 bp mRNA linear INV 10-APR-1998  
 LOCUS Anticardias crassispina mRNA for Csk-type protein tyrosine kinase,  
 DEFINITION partial cds.

ACCESSION AB012713.1 GI:3043753  
 VERSION AB012713.1  
 KEYWORDS Csk-type protein tyrosine kinase; ACNRTK5.  
 SOURCE Anticardias crassispina cDNA to mRNA.  
 ORGANISM Anticardias crassispina  
 Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 Echinoidea; Euechinoidea; Echinacea; Echinoidea; Echinometridae;  
 Anthocidaris.  
 1 (sites)  
 REFERENCE 1 (sites)  
 AUTHORS Sakuma, M., Onodera, H., Suyemitsu, T. and Yamasu, K.  
 TITLE The protein tyrosine kinases of the sea urchin Anthocidaris  
 crassispina  
 JOURNAL Zool. Sci. 14 (6), 941-946 (1997)  
 MEDLINE 98181253  
 REFERENCE 2 (bases 1 to 187)  
 AUTHORS Yamasu, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-MAR-1998) Kyo Yamasu, Saitama University, Faculty of  
 Science, Department of Regulation Biology, 255 Shimo-Ookubo, Urawa,  
 Saitama 338-8570, Japan (E-mail:kyamasu@satl.saitama-u.ac.jp,  
 Tel:+81-48-858-3417, Fax:+81-48-858-3698)  
 FEATURES  
 source  
 1..187  
 /organism="Anticardias crassispina"  
 /db\_xref="taxon:7629"  
 1..187  
 /gene="ACNRTK5"  
 <1..>187  
 /gene="ACNRTK5"  
 /codon\_start=1  
 /product="Csk-type protein tyrosine kinase"  
 /protein\_id="BA25541.1"  
 /db\_xref="GI:3043754"  
 /translation="VHRDLARNILVSGDIAKVSDFGLSGEATLNQEGKFKPIKMTA  
 PEALRRKEFSMSDVWST"  
 BASE COUNT 48 a 44 c 49 g 46 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.0369 Length: 187  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 3 Gaps: 0

US-09-397-967-16 (1-1099) x AB012713 (1-187)

OY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952  
 |||||||||||||||||||||||||||||||||||

Db 1 GTGCACAGAGATCTCGCGCTCGAACAATCCTGCTG 36

RESULT 69  
 HYDTRYKIND  
 LOCUS HYDTRYKIND 201 bp mRNA linear INV 26-APR-1993  
 DEFINITION Hydra vulgaris tyrosine kinase mRNA, partial cds.  
 ACCESSION M64614  
 VERSION M64614.1 GI:159281  
 KEYWORDS tyrosine kinase.  
 SOURCE Hydra vulgaris, cDNA to mRNA.

ORGANISM Hydra vulgaris  
 Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Anthomedusae;  
 Hydridae; Hydra.  
 1 (bases 1 to 201)  
 REFERENCE Steele, R.E.  
 AUTHORS Unpublished (1991)  
 JOURNAL  
 FEATURES  
 source  
 1..201  
 /organism="Hydra vulgaris"  
 /db\_xref="taxon:6087"  
 /dev\_stage="adult polyp"  
 <1..>201  
 /note="putative"  
 /codon\_start=1  
 /protein\_id="AAA29221.1"  
 /db\_xref="GI:159282"  
 /translation="VHRDLARNILVSGAGKVKISDFGLTRKINDELINYSKKRRRLP  
 VKWMSVEAIFDQLTFSFSDVWTF"  
 BASE COUNT 65 a 30 c 47 g 59 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.0391 Length: 201  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 3 Gaps: 0

US-09-397-967-16 (1-1099) x HYDTRYKIND (1-201)

OY 941 VALHISARGASPLEUALAALARGASNIILEUVAL 952  
 |||||||||||||||||||||||||||||||||||

Db 1 GTACATCGGATCTGCTGCAAGAAACATTTTGGTC 36

RESULT 70  
 H006874S11  
 LOCUS H006874S11 206 bp DNA linear PRI 29-JAN-2001  
 DEFINITION Homo sapiens megakaryocyte-associated tyrosine kinase (MATEK) gene,  
 exon 11.  
 ACCESSION S75159  
 VERSION S75159.1 GI:896215  
 KEYWORDS 11 of 13  
 SEGMENT  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 206)  
 REFERENCE 1 (bases 1 to 206)  
 AUTHORS Avraham, S., Jiang, S., Ota, S., Fu, Y., Deng, B., Dowler, L.L.,  
 White, R.A. and Avraham, H.  
 TITLE Structural and functional studies of the intracellular tyrosine  
 kinase MATEK gene and its translated product  
 J. Biol. Chem. 270 (4), 1833-1842 (1995)  
 MEDLINE 95130565  
 PUBMED 7530249  
 REMARK Genbank staff at the National Library of Medicine created this  
 entry [NCBI g1bbsq 159439] from the original journal article.  
 This sequence comes from Fig. 5.  
 FEATURES  
 source  
 1..206  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 6..201  
 /gene="MATEK"  
 /number=11  
 BASE COUNT 43 a 64 c 72 g 27 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.0399 Length: 206  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x H006874511 (1-206)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 |||||  
 Db 49 GTGCACCGGACCTGGCGCCGCGACATCTGCTC 84

RESULT 71  
 AX203083 214 bp DNA linear PAT 30-AUG-2001  
 DEFINITION Sequence 4 from Patent WO0152892.  
 ACCESSION AX203083  
 VERSION AX203083.1 GI:15392440  
 KEYWORDS

SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 214)  
 AUTHORS Vassios,G.  
 TITLE Jak/stat pathway inhibitors and the uses thereof  
 JOURNAL Patent: WO 0152892-A 4 26-JUL-2001;  
 GENZYME CORPORATION (US)

FEATURES  
 source Location/Qualifiers  
 1..214  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 34 a 74 c 62 g 36 t 8 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.0411 Length: 214  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AX203083 (1-214)

QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsn 949  
 |||||  
 Db 128 CCGCGCTGCGTGCACCGCGACCTGGCGCCGCGAANC 163

RESULT 72  
 AB038987 302 bp mRNA linear VRT 28-AUG-2001  
 LOCUS Gallus gallus EphA9 mRNA for protein tyrosine kinase EphA9, partial  
 DEFINITION cds.  
 ACCESSION AB038987  
 VERSION AB038987.1 GI:15375362  
 KEYWORDS protein tyrosine kinase EphA9,  
 SOURCE Gallus gallus Stage X embryos Blastodermal Cells cDNA to mRNA.  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.

REFERENCE 1  
 Sasaki,E., Sakurai,M. and Etches,R.J.  
 TITLE Expression of Protein Tyrosine Kinase in Chicken Blastodermal Cells  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 302)  
 AUTHORS Sasaki,E., Sakurai,M. and Etches,R.J.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-FEB-2000) Erika Sasaki, University of Guelph,  
 Department of Food Science, Guelph, Ontario, N1G 2W1, Canada  
 (E-mail:enakaminuoguelph.ca, Tel:1-519-824-4120(ex.2259),  
 Fax:1-519-824-6631)  
 FEATURES  
 source Location/Qualifiers  
 1..302

gene  
 CDS  
 1..302  
 /gene="EphA9"  
 <1..>302  
 /gene="EphA9"  
 /codon\_start=3  
 /product="protein tyrosine kinase EphA9"  
 /protein\_id="BAB63911.1"  
 /db\_xref="GI:15375363"  
 /translation="NGALDTFLRENEERFSPVQVLYNMLQGLASGNTYLSHHNYVHRL  
 AARNILVTRSLQCKVSDPFLSRITLNDAGTYETKGIPIRWTAPEAIARHIFTS"

BASE COUNT 80 a 75 c 76 g 71 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.0542 Length: 302  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x AB038987 (1-302)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 |||||  
 Db 120 GTGCACCGGACCTGGCTGCTCGACATCTGTGTA 155

RESULT 73  
 CHRTYKDA 373 bp DNA linear VRT 28-APR-1993  
 LOCUS CHRTYKDA  
 DEFINITION Chicken DNA that hybridizes to v-fps tyrosine kinase domain of the  
 Rous sarcoma virus.  
 ACCESSION M12257  
 VERSION M12257.1 GI:212845  
 KEYWORDS  
 SOURCE Chicken DNA, clone pTKR11.  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 373)  
 POSTER,F.D.A., Levy,J.B., Daley,G.O., Simon,M.C. and Hanafusa,H.  
 TITLE Isolation of chicken cellular DNA sequences with homology to the  
 region of viral oncogenes that encodes the tyrosine kinase domain  
 JOURNAL Mol. Cell. Biol. 6 (1), 325-331 (1986)  
 MEDLINE 87064304  
 PUBMED 3023834

COMMENT Draft entry and computer readable copy of sequence [1] kindly  
 provided by D.A.Foster, 30-MAY-1986.  
 FEATURES  
 source Location/Qualifiers  
 1..373  
 /organism="Gallus gallus"  
 /db\_xref="taxon:9031"

BASE COUNT 90 a 90 c 104 g 89 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.0642 Length: 373  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x CHRTYKDA (1-373)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 |||||  
 Db 60 GTGCACCGGACCTGGCTGCGCCGCGACATCTGCTC 95

```

RESULT 74
AR053289      738 bp      DNA      linear      PAT 29-SEP-1999
LOCUS         AR053289
DEFINITION    Sequence 8 from patent US 5834206.
ACCESSION     AR053289
VERSION       AR053289.1 GI:5978151
KEYWORDS
SOURCE        Unknown.
ORGANISM      Unknown.
REFERENCE     1 (bases 1 to 738)
AUTHORS      Sakano,S.
TITLE        Tyrosine kinase
JOURNAL      Patent: US 5834208-A 8 10-NOV-1998;
FEATURES
source       1..738
              /organism="unknown"
BASE COUNT   154 a 205 c 248 g 131 t
ORIGIN
Alignment Scores:
Pred. No.:    0.111      Length:      738
Score:        12.00      Matches:     12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:  1.09%      Indels:      0
DB:           6          Gaps:          0
US-09-397-967-16 (1-1099) x AR053289 (1-738)
QY  941  VALHSAARGASPLEUALAALARGASNLLEUVAL 952
Db  349  GTGACCGCGACCTGGCCGCCGACATCTGTGTC 384

RESULT 75
RNO299010     790 bp      DNA      linear      ROD 05-MAR-2002
LOCUS         RNO299010
DEFINITION    Rattus norvegicus partial Ret gene for receptor tyrosine kinase,
              exons 14-15.
ACCESSION     AJ299010
VERSION       AJ299010.1 GI:10638054
KEYWORDS      receptor tyrosine kinase; ret gene; RET proto-oncogene.
SOURCE        Norway rat.
ORGANISM      Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE     1
AUTHORS      Matera,I., De Miguel-Rodriguez,M., Fernandez-Santos,J.M.,
              Santamaria,G., Puliti,A., Ravazzolo,R., Romeo,G.,
              Galera-Davidson,H. and Ceccherini,I.
TITLE        DNA sequence and genomic structure of the rat RET proto-oncogene
JOURNAL      Mol. Cell. Neurosci. 11 (5), 405-417 (2000)
MEDLINE       21226279
PUBMED        11328649
REFERENCE     2 (bases 1 to 790)
AUTHORS      Ceccherini,I.
TITLE        Direct Submission
JOURNAL      Submitted (21-SEP-2000) Ceccherini I., Laboratorio Genetica
              Molecolare, Istituto Glannina Gaslini, Largo G. Gaslini, 5 -
              Genova, 16148, ITALY
FEATURES
source       1..790
              /organism="Rattus norvegicus"
              /db_xref="taxon:10116"
              1..790
              /gene="Ret"
              <1..283
              /gene="Ret"
              /number=13
              284..498
              /gene="Ret"

```

```

/number=14
/usedin-AJ298999:Ret_CDS1
/usedin-AJ298999:Ret_CDS2
/usedin-AJ298999:Ret_mRNA1
/usedin-AJ298999:Ret_mRNA2
499..591
/gene="Ret"
/number=14
592..714
/gene="Ret"
/number=15
/usedin-AJ298999:Ret_CDS1
/usedin-AJ298999:Ret_CDS2
/usedin-AJ298999:Ret_mRNA1
/usedin-AJ298999:Ret_mRNA2
715..>790
/gene="Ret"
/number=15
Intron
BASE COUNT   178 a 203 c 230 g 177 t 2 others
ORIGIN
Alignment Scores:
Pred. No.:    0.117      Length:      790
Score:        12.00      Matches:     12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:  1.09%      Indels:      0
DB:           10          Gaps:          0
US-09-397-967-16 (1-1099) x RNO299010 (1-790)
QY  941  VALHSAARGASPLEUALAALARGASNLLEUVAL 952
Db  595  GTACATCGAGACTTACTGCCAGAACATCTTGTC 630

RESULT 76
DR089380     877 bp      mRNA      linear      VRT 26-JUN-1997
LOCUS         DR089380
DEFINITION    Danio rerio Eph-like kinase 2 (zek2) mRNA, partial cds.
ACCESSION     U89380
VERSION       U89380.1 GI:2198796
KEYWORDS
SOURCE        Danio rerio.
ORGANISM      Danio rerio
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
              Cypriniformes; Cyprinidae; Danio.
REFERENCE     1 (bases 1 to 877)
AUTHORS      Bovenkamp,D.E. and Greer,P.
TITLE        Novel Eph-family receptor tyrosine kinase is widely expressed in
              the developing zebrafish nervous system
JOURNAL      Dev. Dyn. 209 (2), 166-181 (1997)
MEDLINE       9186052
PUBMED        9186052
REFERENCE     2 (bases 1 to 877)
AUTHORS      Bovenkamp,D.E. and Greer,P.
TITLE        Direct Submission
JOURNAL      Submitted (13-FEB-1997) Department of Biochemistry, Cancer Research
              Labs, Queen's University, Rm. A309 Botterell Hall, Kingston, ON K7L
              3N6, Canada
FEATURES
source       1..877
              /organism="Danio rerio"
              /db_xref="taxon:7955"
              <1..>877
              /gene="zek2"
              <1..>877
              /gene="zek2"
              /note="Eph-like kinase 2; zek2; Eph-like receptor tyrosine
              kinase"
              /codon_start=1
              /product="zek2"
              /protein_id="AAC60222.1"

```

```

/misc-feature
1. .828

```

BASE COUNT		245 a	326 c	308 g	249 t	/gene="Sek-4"
ORIGIN						/note="intra cellular region"
Alignment Scores:						
Pred. No.:	0.155	Length:	1128			
Score:	12.00	Matches:	12			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	1.09%	Indels:	0			
DB:	10	Gaps:	0			
US-09-397-967-16 (1-1099) x MMSEK4 (1-1128)						
Oy	941	ValHisArgAspLeuAlaAlaArgAsnLeuVal 952				
Db	94	GTGCACCGTGACCTTGCTGCCCGAACAATCCTTCTC 129				
RESULT 78						
LOCUS	AB025542	1153 bp	mRNA	linear	VRT 18-NOV-1999	
DEFINITION	Lampetra reissneri mRNA for Ephb, partial cds.					
ACCESSION	AB025542					
VERSION	AB025542.1	GI:6002426				
KEYWORDS	Ephb.					
SOURCE	Lampetra reissneri larva cDNA to mRNA.					
ORGANISM	Lethenteron reissneri					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia; Petromyzontiformes; Petromyzontidae; Lethenteron.					
AUTHORS	1 (siles)					
TITLE	Suga,H., Hoshiyama,D., Kuraku,S., Katoh,K., Kubokawa,K. and Miyata,T.					
JOURNAL	Protein tyrosine kinase cDNAs from amphioxus, hagfish, and lamprey: isoform duplications around the divergence of cyclostomes and gnathostomes					
REFERENCE	J. Mol. Evol. 49 (5), 601-608 (1999)					
MEDLINE	20020330					
AUTHORS	2 (bases 1 to 1153)					
TITLE	Suga,H.					
JOURNAL	Direct Submission					
REFERENCE	Submitted (31-MAR-1999) Hiroshi Suga, Faculty of Science, Kyoto University, Department of Biophysics, Kitashirakawa-oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan					
FEATURES	(E-mail:sugah@biophys.kyoto-u.ac.jp, Tel:+81-75-753-4224, Fax:+81-75-753-4223)					
source	Location/Qualifiers					
CDs	1..1153					
	/organism="Lethenteron reissneri"					
	/db_xref="taxon:7753"					
	/dev_stage="larva"					
	<1..1063					
	/codon_start=2					
	/product="Ephb"					
	/protein_id="BAA84732.1"					
	/db_xref="GI:6002427"					
	/translation="EVCSGRLLKPGKREMAVAIKTLKAGYTERORPDFLSASTMGQF					
	DHPVTHLEGVYTKSPVMVITTEFMENSLDSFLRNDGFTVQLVGMGLGIAAGMK					
	YVADKSVYRDLAATVNLVNSLVCKVDFGSLRLEDGTDSPYTSAMGKILIRMT					
	AFALQIKRFTTASDVWSTGIVIMEVMSVGEPPWDSNNOVIAIEDYRLPAPMCO					
	PAALQILMDCKQKERNRPRKGOIVNLTDLIRPAISLKVACTPDISPLDRSDV					
	PPGFASGVSDVEMIDAIKMKRVDNFRASGFTISFDLVSMVTADILIRICVTLIAGHKKK					
	LNISIQTRVQNMQIQTVQY"					
BASE COUNT	275 a	330 c	335 g	213 t		
ORIGIN						
Alignment Scores:						
Pred. No.:	0.158	Length:	1153			
Score:	12.00	Matches:	12			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	1.09%	Indels:	0			
DB:	5	Gaps:	0			



US-09-397-967-16 (1-1099) x AB025542 (1-1153)

Oy 941 ValHISARGSPLEUALAALARGASNIleleVal 952  
Db 329 GTGCATCGAGATCTGCGACGACGACATCTGTC 364

RESULT 79  
LOCUS HYDTRYRKINA 1186 bp mRNA linear INV 08-JUL-1999  
DEFINITION Hydra vulgaris putative tyrosine kinase mRNA, complete cds.  
ACCESSION M64611  
VERSION M64611.1 GI:159275  
KEYWORDS  
SOURCE Hydra vulgaris.  
ORGANISM Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Anthomedusae;  
Hydridae; Hydra.

REFERENCE 1 (bases 1 to 1186)  
AUTHORS Krolher, M., Reidling, J.C. and Steele, R.E.  
TITLE A gene whose major transcript encodes only the substrate binding domain of a protein-tyrosine kinase

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1186)  
AUTHORS Steele, R.E.

REFERENCE 3 (bases 1 to 1186)  
AUTHORS Krolher, M., Reidling, J.C. and Steele, R.E.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAY-1991) Department of Biological Chemistry, University of California, 240D Medical Sciences I, Irvine, CA 92697-1700, USA

FEATURES  
source Location/Qualifiers  
1..1186

/organism="Hydra vulgaris"  
/db\_xref="taxon:6087"  
/dev\_stage="adult polyp"  
110..823  
/codon\_start=1  
/product="putative tyrosine kinase"  
/protein\_id="AA029219.1"  
/db\_xref="GI:159276"

/translation="MIGCSTLQKPICLVEMENDLQFLKNSRTKLITSLYKSS  
VFENYQMPPEYLSIKILVHRLAARNILGADKIYKISDFGLTRVNDIDYNGS  
NNRRLPIKMSYEATFEDRTFTSYDVMAVGIYLPITVGLGSPYSISNRELPLKKS  
CYRMERPNCSMDFCMNCNENEDPLQRPFTTKLRETFEELMSGSNTLNEISEEC  
TYINWPSHSHNSNDLVV"

BASE COUNT 411 a 154 c 196 g 425 t  
ORIGIN

#### Alignment Scores:

Pred. No.: 0.162 Length: 1186  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 3 Gaps: 0

US-09-397-967-16 (1-1099) x HYDTRYRKINA (1-1186)

Oy 941 ValHISARGSPLEUALAALARGASNIleleVal 952  
Db 302 GTACATCGAGATCTGCTCTAGAAATATTATTAGTA 337

#### RESULT 80

LOCUS HUMERK 1225 bp mRNA linear PRI 23-JUN-1999  
DEFINITION Human mRNA for large erk kinase.  
ACCESSION D14717  
VERSION D14717.1 GI:285916  
KEYWORDS large erk kinase.  
SOURCE Homo sapiens gastric carcinoma, CDNA to mRNA, clone H1.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Chan, J. and Watt, V.M.  
TITLE eek and erk, new members of the eph subclass of receptor protein-tyrosine kinases  
JOURNAL Oncogene 6 (6), 1057-1061 (1991)  
MEDLINE 91296384

REFERENCE 2 (sites)  
AUTHORS Iwase, T., Tanaka, M., Suzuki, M., Naito, Y., Sugimura, H. and Kino, I.  
TITLE Identification of protein-tyrosine kinase genes preferentially expressed in embryo stomach and gastric cancer  
JOURNAL Biochem. Biophys. Res. Commun. 194 (2), 698-705 (1993)  
MEDLINE 93343925

REFERENCE 3 (bases 1 to 1225)  
AUTHORS Iwase, T.  
JOURNAL Unpublished  
COMMENT Submitted (23-MAR-1993) to DDBJ by:  
Toshio Iwase  
First Department of Pathology  
Hamamatsu University  
School of Medicine  
3600 Handa-cho, Hamamatsu  
Shizuoka 431-31  
Japan

Phone: 053-435-2220  
Email: toiwase@dbj.nig.ac.jp  
Fax: 053-435-2225

#### FEATURES

source Location/Qualifiers  
1..1225  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="gastric carcinoma"

1..1225  
/gene="large erk"

<1..1044  
/gene="large erk"  
/function="protein-tyrosine kinase"

/codon\_start=1  
/product="large erk kinase"  
/protein\_id="BA03537.1"

/db\_xref="GI:285917"

/translation="LKDPKREITVAIKTKSGYTEKQRDFLSERSIMGQFDHPNVI  
HLEGVTKSTVMTTEPMENGLDSELRDQGFVYIQVGLMGLIAGKRYLADNN  
YVHDLARNILVSNLVCKVSDGLSFLDDYSDPTYSALGCKIPRMTAPALQ  
YRKFTSADVSYSYIVMVEVSYGERPMDTNDGVINALIODYRLPPMDCPALHO  
LMLDCWQDRNHRPRKFGQIVNTLDKMRNPNSLKAMPLSSGINTPLDRTIPYTSF  
NTVDENLKAIKMGQYKESFANAGFTSDVYSQMMEDILRVGTVLAGHOKKILNSIOV  
MRAQMDIQSYEV"

unsure

94..95  
/gene="large erk"  
/citation=[1]  
/replace="gc"  
220..225  
/gene="large erk"  
/citation=[1]  
/replace="gtagg"

BASE COUNT 283 a 386 c 329 g 227 t  
ORIGIN

#### Alignment Scores:

Pred. No.: 0.166 Length: 1225  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x HUMERK (1-1225)

Oy 941 ValHISARGSPLEUALAALARGASNIleleVal 952  
Db 310 GTTCACCGTAGACTGTGCCCGGACACATCCCTGTC 345

RESULT 81  
MMSEK2  
LOCUS M.musculus Sek-2 mRNA. 1294 bp mRNA linear ROD 10-MAR-2001  
DEFINITION X76010.X57343  
ACCESSION X76010.1 GI:460053  
VERSION  
KEYWORDS Sek gene.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1294)  
AUTHORS Becker, N., Seitanidou, T., Murphy, P., Mattei, M.G., Topilko, P., Nieto, M.A., Wilkinson, D.G., Charnay, P. and Giliardi-Hedenstreit, P.  
TITLE Several receptor tyrosine kinase genes of the Eph family are segmentally expressed in the developing hindbrain  
JOURNAL Mech. Dev. 47 (1), 3-17 (1994)  
MEDLINE 95034306  
PUBMED 7947319  
REFERENCE 2 (bases 1 to 1294)  
AUTHORS Becker, N.  
TITLE Direct Submission  
JOURNAL Submitted (11-OCT-1993) N. Becker, Unite Inserm U368 Ecole Normale, Supérieure, Laboratoire de Biologie Moléculaire, du Développement, Patrick Charnay, U368 INSERM 46 Rue d'Ulm 75005 Paris, FRANCE  
FEATURES  
SOURCE  
1. 1294  
/organism="Mus musculus"  
/strain="C57 BL"  
/db\_xref="taxon:10090"  
/chromosome="4"  
/map="D.E"  
/clone="1"  
/tissue-type="total embryo"  
/clone\_lib="Brightle HOGAN"  
/dev\_stage="embryo, 8.5 days"  
1. 1281  
/gene="Sek-2"  
/locus="1.1281"  
/gene="Sek-2"  
/codon\_start=1  
/protein\_id="CAA53597.1"  
/db\_xref="GI:460054"  
/db\_xref="SWISS-PROT:Q03145"  
/translation="LAGGLFIRRRRLRAROSSEVDYRFSEKSLKPLKTYVDPHTY  
EDPNAVLTFTTEIHPSCVAKOVYGAEEFGEVKGTLKASSGKEIPVALIKTLKAGY  
TEKORVDFLSASITMGPSHHNIIIRLEGVSKYKPMITTEMGALDKFLREKDE  
FSVCLVLMKGLASGMKYLNNMTVHDLARNILVSNLYCVSPDGLSRVLEDDP  
EAYVTTSGSKLPIRWTAPEALSYRFTASDVSTGIVMEVMTYGERPYWELSNHEY  
MKAINDGRLPTMDGPSAIVQLMMQOQOERSRPRPADIVSLDKLIRAPDSIKTL  
ADFPVRVIRLPTSGSEGVPRVSEMISEIKMOQYEHFMVAGYTAIEVYVOMSNR  
DIKRGVRLPGHOKRIAYSLGLKQVNTVGIPI"  
1. 27  
/gene="Sek-2"  
/note="Trans membrane region"  
misc\_feature  
/note="Intra cellular region"  
misc\_feature  
/gene="Sek-2"  
28. 1281  
/note="Intra cellular region"  
BASE COUNT 314 a 363 c 365 g 252 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.174 Length: 1294  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0  
DB: 10  
US-09-397-967-16 (1-1099) x MMSEK2 (1-1294)  
Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 556 GTGCACCGGAGCTGGCCGCCGACATCCTGTC 591  
RESULT 82  
AR053290  
LOCUS AR053290 1398 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 9 from patent US 5834208.  
ACCESSION AR053290  
VERSION AR053290.1 GI:5978152  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1398)  
AUTHORS Sakano, S.  
TITLE Tyrosine kinase  
JOURNAL Patent: US 5834208-A 9 10-NOV-1998;  
FEATURES  
source  
1. 1398  
/organism="unknown"  
BASE COUNT 282 a 426 c 468 g 222 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.185 Length: 1398  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0  
US-09-397-967-16 (1-1099) x AR053290 (1-1398)  
Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
Db 922 GTGCACCGGAGCTGGCCGCCGACATCCTGTC 957  
RESULT 83  
DRAJ5028  
LOCUS DRAJ5028 1473 bp mRNA linear VRT 27-MAY-1998  
DEFINITION Danio rerio mRNA for Eph-like receptor tyrosine kinase rtk7,  
partial.  
ACCESSION AJ005028  
VERSION AJ005028.1 GI:3005904  
KEYWORDS Eph-like receptor tyrosine kinase.  
SOURCE zebrafish.  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.  
REFERENCE 1 (bases 1 to 1473)  
AUTHORS Cooke, J.E.  
TITLE Direct Submission  
JOURNAL Submitted (27-MAR-1998) Cooke J.E., University College London, Department of Anatomy and Developmental Biology, Gower Street, London, WC1E 6BT, UK  
2 (bases 1 to 1473)  
Cooke, J.E., Xu, Q., Wilson, S.W. and Holder N.  
Characterisation of five novel zebrafish Eph-related receptor tyrosine kinases suggests roles in patterning the neural plate  
Dev. Genes Evol. 206, 515-531 (1997)  
FEATURES  
source  
1. 1473  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/db\_xref="taxon:7955"  
/codon\_start=1  
/product="Eph-like receptor tyrosine kinase rtk7"  
/protein\_id="CAA06301.1"  
/db\_xref="GI:3005905"  
/db\_xref="SPTREMBL:O73877"  
/translation="KSNMTVDGKPGTTFYFRVARPDGQYGVKGEILETSHEDM  
LAVGPNOQTILATISVAGAVLVLVACFIVSGRCGYIKAKODPEEKMOFGHRY

KLPTRTYIHPTIEDPNOAVDFAKELIENSIRIERYGAGEFEGECGRLRLPSKR  
 EIOVAIKSLKAGYSEHORRDLSEASIMGQFDHPNIIIEGVTRCKPMIYTEMEN  
 GSDPTFKHDSGOFTVIOILGMOLYLSPMNVHDLARNILVNGVNCYK  
 SDPGISRLIEDDEPEAYTTRGSKIPIRMTAPAEATYTRKETSASDMSYGIYMEVYSY  
 GERPMEMSNDVYKAVDEGRRLPRMPCPYVYLHOLMDCEKNSDPRKQIYNTL  
 DRLIRNPSLKLQANSVAVDEVPVEAVNVTEMDLIRKGOYKREHPSAGYVTLDS  
 VLYVSSSLDMLGVELAGHOKRKLISLIDCAHHGTQYOV"

BASE COUNT 387 a 341 c 431 g 314 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 0.192 Length: 1473  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x DRAJ5028 (1-1473)

OY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952  
 DB 769 GTGCATCGGAGATTGTGCACGACGAACTTCGTGTA 804

## RESULT 84

RNELK 1509 bp mRNA linear ROD 14-JUL-1995  
 LOCUS RNELK  
 DEFINITION Rat mRNA for elk protein.  
 ACCESSION X13411  
 VERSION X13411.1 GI:56094  
 KEYWORDS elk gene; protein-tyrosine kinase.  
 SOURCE Rattus rattus  
 ORGANISM Rattus rattus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 1509)  
 AUTHORS Letwin,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-NOV-1988) Letwin K., Mount Sinai Hospital Research  
 Institute, 600 University Avenue, Toronto, Ontario, M5G 1X5  
 REFERENCE 2 (bases 1 to 1509)  
 AUTHORS Letwin,K., Yee,S.P. and Pawson,T.  
 TITLE Novel protein-tyrosine kinase cDNAs related to fps/fee and eph  
 cloned using anti-phosphotyrosine antibody  
 JOURNAL Oncogene 3 (6), 621-627 (1988)  
 MEDLINE 94167102  
 PUBMED 2485255  
 COMMENT Data kindly reviewed (03-Feb-1989) by Letwin K.  
 FEATURES  
 SOURCE  
 location/Qualifiers  
 1..1509  
 /organism="Rattus rattus"  
 /strain="Wistar"  
 /db\_xref="taxon:10117"  
 /clone="lambda B1"  
 /tissue\_type="brain"  
 /dev\_stage="adult"  
 <1..1143  
 /codon\_start=1  
 /product="elk protein"  
 /protein\_18="CA31777.1"  
 /db\_xref="GI:56095"  
 /db\_xref="SWISS-PROT:P09759"  
 /translation="EAVREFAKEIDVSPKITEVIGAGEFGEVYKGRLLPGKREIY  
 AITKLKAGYSEHORRDLSEASIMGQFDHPNIIIEGVTRCKPMIYTEMEN  
 SELNDGQFTVIOILGMOLYLSPMNVHDLARNILVNGVNCYKSDPG  
 LSRYLDPTSDPTSSLSGKIPYRMTAPAEATYTRKETSASDMSYGIYMEVYSFGE  
 RPYDMSQDVINAIEDYRLPPEDCPAALHQLMDCWQDRNSRPRFAIVNTLCK  
 MIRNPASLKYATITVAVPSQPLDRSLPDTFAITVDMLSAIKWQYRDSFLTAGFT  
 SLQLVQMTSEDLIRIGVTLAGHOKRKLISSISHKRVQNSPSVMA"  
 23..274  
 /note="catalytic domain"

misc\_feature  
 BASE COUNT 393 a 373 c 430 g 313 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 0.196 Length: 1509  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x RNELK (1-1509)

OY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952  
 DB 409 GTGCACCGGACCTGCGCTCTAGAACATTCGTG 444

## RESULT 85

AR053291 1521 bp DNA linear PAT 29-SEP-1999  
 LOCUS AR053291  
 DEFINITION Sequence 10 from patent US 5834208.  
 ACCESSION AR053291  
 VERSION AR053291.1 GI:5978153  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1521)  
 AUTHORS Sakano,S.  
 TITLE Tyrosine kinase  
 JOURNAL Patent: US 5834208-A 10-10-NOV-1998;  
 FEATURES  
 source  
 location/Qualifiers  
 1..1521  
 /organism="unknown"

BASE COUNT 295 a 470 c 507 g 249 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0.197 Length: 1521  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 6 Gaps: 0

US-09-397-967-16 (1-1099) x AR053291 (1-1521)

OY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952  
 DB 1045 GTGCACCGGACCTGCGCGCCGCAACATTCGTGTC 1080

RESULT 86  
 LOCUS XLEPHRTK 1531 bp mRNA linear VRT 02-APR-1996  
 DEFINITION Xenopus laevis eph receptor tyrosine kinase subfamily mRNA, 3' end  
 of cds.  
 ACCESSION L43622  
 VERSION L43622.1 GI:974707  
 KEYWORDS eph-class receptor tyrosine kinase; receptor tyrosine kinase.  
 SOURCE Xenopus laevis (clone: PL7a) gastrula cDNA to mRNA.  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;

REFERENCE 1 (bases 1 to 1531)  
 AUTHORS Scales,J.B., Wining,R.S., Renaud,C.S., Shea,L.J. and Sargent,T.D.  
 TITLE Novel members of the eph receptor tyrosine kinase subfamily  
 expressed during Xenopus development  
 JOURNAL Oncogene 11 (9), 1745-1752 (1995)  
 MEDLINE 96068901  
 PUBMED 7478602  
 FEATURES  
 source  
 location/Qualifiers  
 1..1531  
 /organism="Xenopus laevis"

/db\_xref="taxon:8355"  
 /clone="pu7a"  
 /dev\_stage="gastrula"  
 <1..972  
 /codon\_start=1  
 /product="Eph receptor tyrosine kinase"  
 /protein\_id="AA093525.1"  
 /db\_xref="GI:974708"  
 /translation="EFISFASIMGQFNHPIHLEGVITNNGPVMTTEYMENGALDS  
 FLKNDQGFPIQVGLKGIASGRIYSEMSVTHDLARNILVNSNLCKVSDGEL  
 SRFLOGSTDPYITSLGKIPIKWTAPKAIKFTASDVWGIVGIVMEVSEFGR  
 PYWMSNOVDYINAEODYRLPAPDCPTALHQLMDCOWDRALRFRFADIVSLDKL  
 IRNPASLKITTRFOLSTQPLDORTPHYSSFSVSEMLHAIKMGREYEDGFRNAGFTT  
 FSRVONISTEDILRMGVTLAGHOKIILSSIQILPSEKSTPIPDHCY"  
 BASE COUNT 377 a 410 c 383 g 361 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.199 Length: 1531  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 5 Gaps: 0  
 US-09-397-967-16 (1-1099) x XELEPHRTK (1-1531)  
 QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 Db 232 GTGCACCGGAGATCTGCGCCAGAAATATCTAGTG 267  
 RESULT 87  
 LOCUS AB025541 1563 bp mRNA linear VRT 18-NOV-1999  
 DEFINITION Eptaretus burgeri mRNA for EphA, partial cds.  
 ACCESSION AB025541  
 VERSION AB025541.1 GI:6002424  
 KEYWORDS EphA.  
 SOURCE Eptaretus burgeri adult female cDNA to mRNA.  
 ORGANISM Eptaretus burgeri  
 Eukaryota; Metazoa; Chordata; Craniata; Hyperotretli; Myxiniiformes;  
 Myxiniidae; Eptaretinae; Eptaretus.  
 1 (sites)  
 Suga,H., Hoshiyama,D., Kuraku,S., Katch,K., Kubokawa,K. and  
 Miyata,T.  
 Protein tyrosine kinase cDNAs from amphioxus, hagfish, and lamprey:  
 isoform duplications around the divergence of cyclostomes and  
 gnathostomes  
 J. Mol. Evol. 49 (5), 601-608 (1999)  
 2 (bases 1 to 1563)  
 Suga,H.  
 Direct Submission  
 Submitted (31-MAR-1999) Hiroshi Suga, Faculty of Science, Kyoto  
 University, Department of Biophysics, Kitashirakawa-oiwakecho,  
 Sakyo-ku, Kyoto 606-8502, Japan  
 (E-mail:suga@biophys.kyoto-u.ac.jp, Tel:+81-75-753-4224,  
 Fax:+81-75-753-4223)  
 FEATURES  
 source  
 1.1563  
 location/Qualifiers  
 /organism="Eptaretus burgeri"  
 /db\_xref="taxon:7764"  
 /sex="female"  
 /dev\_stage="adult"  
 <1..1063  
 /codon\_start=2  
 /product="EphA"  
 /protein\_id="BAA84731.1"  
 /db\_xref="GI:6002425"  
 /translation="EVCSGYLKLGRREVCVAIKTKAGTYEQORDFLCASIMGCF  
 DHEINIIMEGVITKRSKPVMTITPEFMENGLDLFLRKNDGQFTAMOLVGMRLGASGMR  
 YLADMSYVHRDLAARNILVNGNIVCVKVSDEGLSRVLEDDDAAYMRGCKIPKMTAP  
 EALAFKRTSASDVMSYGIYVMEVYTSIGERPYWMSNOVDYIKAEISGYRLPAPDCPL

ALHQLMDCORERAVRPKFGHIVTILDKIRNPATLKPANEGTMRVPGSLDQSTD  
 NPAVRTVPEMLAEIKMGRYADNNAAGYTTLDSVAQLDGLVIRIGITILGHOKKILS  
 NIOTMKAEMLDHSRGVH"  
 BASE COUNT 389 a 371 c 421 g 382 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.202 Length: 1563  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 5 Gaps: 0  
 US-09-397-967-16 (1-1099) x AB025541 (1-1563)  
 QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 Db 329 GTACACCGTGACCTTGCTGCCAGAAATATCTGCTT 364  
 RESULT 88  
 LOCUS HSURFRET 1568 bp mRNA linear PRI 14-JAN-1992  
 DEFINITION H.sapiens urf-ret mRNA.  
 ACCESSION X56348  
 VERSION X56348.1 GI:37611  
 KEYWORDS urf-ret mRNA.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 1568)  
 Kunieda,T.  
 Direct Submission  
 Submitted (03-SEP-1990) T. Kunieda, IMAMICHI INSTITUTE FOR ANIMAL  
 REPRODUCTION, FUKAYA 1103, DEJIMA-MURA, NIINHARI-GUN, IBARAKI  
 300-01, JAPAN  
 2 (bases 1 to 1568)  
 Kunieda,T., Matsui,M., Nomura,N. and Ishizaki,R.  
 Cloning of an activated human ret gene with a novel 5' sequence  
 fused by DNA rearrangement  
 JOURNAL Gene 107 (2), 323-328 (1991)  
 MEDLINE 92084126  
 PUBMED 1748302  
 FEATURES  
 source  
 1.1568  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /isolate="H116"  
 /db\_xref="taxon:9606"  
 /cell\_line="NIH3T3 cells transformed by human DNA"  
 /dev\_stage="stomach cancer"  
 194..1501  
 /gene="urf-ret"  
 194..1501  
 /gene="urf-ret"  
 /codon\_start=1  
 /protein\_id="CAA39792.1"  
 /db\_xref="GI:37612"  
 /translation="MEELIYELRLFLFELDHEVLTSTVRKKAVITNILLRIQSSGCF  
 DYVDNAQOKQTANSLPAPQMPPEIPOPMLDPPKKEFPKKNIVLKTIGEGEFGYV  
 KATAFHLKGRAGTIVAVKMLKENASPELRLDLSEFNVLKOVNHHVYIKLYACSD  
 GATLLIVKAYKSLGFLRESKRVGPYLGSGSRRSSLDHPDERALTMGLISFA  
 MQISQMOYLAEMKLVHRLAARNILVAERKKKISDFGLSRDYEDSVYKRSQRI  
 PVKMAIESLFDHIYITGSDVMSFGVLAELIYVILGPNYPGILPERLFINLKGHME  
 RPDNCSEEMVRLMLQCKOPEKRPVFAVISKDLKEMVKKRQYLDLAASPDSILY  
 DGLSDEEPYIVCNNAPLPRAIPSPWIKLYGRISHAFTR"  
 BASE COUNT 359 a 430 c 451 g 328 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.202 Length: 1568  
 Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x HSURERET (1-1568)

OY 941 VALHISARGASPLEUVALAALARGASNIILEUVAL 952  
DB 893 GTTCATCGGACACTTGGCCAGCAACATCTCTGTA 928

RESULT 89  
AB025538  
LOCUS Eptatretus burgeri mRNA for Ephb, partial cds. 1582 bp mRNA linear VRT 18-NOV-1999  
ACCESSION AB025538.1 GI:6002418  
VERSION AB025538.1  
KEYWORDS Ephb.  
SOURCE Eptatretus burgeri adult female cDNA to mRNA.  
ORGANISM Eukaryota; Metazoa; Chordata; Hyperotreti; Myxiniiformes; Myxiniidae; Eptatretinae; Eptatretus.  
REFERENCE 1 (sites)  
AUTHORS Suga, H., Hoshlyama, D., Kuraku, S., Katoh, K., Kubokawa, K. and Miyata, T.  
TITLE Protein tyrosine kinase cDNAs from amphioxus, hagfish, and lamprey: isoform duplications around the divergence of cyclostomes and gnathostomes  
JOURNAL J. Mol. Evol. 49 (5), 601-608 (1999)  
MEDLINE 20020330  
REFERENCE 2 (bases 1 to 1582)  
AUTHORS Suga, H.  
TITLE Direct Submission  
JOURNAL Submitted (31-MAR-1999) Hiroshi Suga, Faculty of Science, Kyoto University, Department of Biophysics, Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan  
FEATURES  
source location/Qualifiers  
1..1582  
/organism="Eptatretus burgeri"  
/db\_xref="taxon:7764"  
/sex="female"  
/dev\_stage="adult"  
<1..1063  
/codon\_start=2  
/product="Ephb"  
/protein\_id="BA84728.1"  
/db\_xref="GI:6002419"  
/translation="EVCSGRLLKPKREMSVAITLAKGYTDRODRDPLSEASTMGOF  
DHPNVILHEGVYTKSRPYMITTEEMENGSIDSLFRONDGOPTVTOIGMLRGIAAGK  
YLAMSVYHRLAARNLIVNSLIVKCYSDSLSELEDDTSPTITSAMGGKIPINVT  
APEAIQYKKTPTASDWSYCYIVMEVMSYGERPYWDSNDQVNAIDQYRLPAPMOC  
PALHQLMLDSWOKERARERKRGIVTLDLITTPASLSLTVAACVTGIGISOPLDRLT  
PVTSTFATVDEMDALIKRGYKKNFANAGSSFDVVSRTMAEDILRIIGVILAGHOKKI  
LNSIQTMVQNOIOIYOV"

BASE COUNT 422 a 348 c 427 g 385 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.204 Length: 1582  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x AB025538 (1-1582)

OY 941 VALHISARGASPLEUVALAALARGASNIILEUVAL 952  
DB 329 GTGCATCGACATCTGCTGCCAGAAACATCTTGGTG 364

RESULT 90  
AB049595  
LOCUS Ephydatia fluviatilis EFPTK150 mRNA for protein tyrosine kinase. 1585 bp mRNA linear INV 27-DEC-2001  
ACCESSION AB049595  
VERSION AB049595.1 GI:18146651  
KEYWORDS Ephydatia fluviatilis cDNA to mRNA.  
SOURCE Ephydatia fluviatilis  
ORGANISM Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha; Haplosclerida; Spongiillidae; Ephydatia.  
REFERENCE 1  
AUTHORS Suga, H., Katoh, K. and Miyata, T.  
TITLE Sponge homologs of vertebrate protein tyrosine kinases and frequent domain shufflings in the early evolution of animals before the parazoan-eumetazoan split  
JOURNAL Gene 280 (1-2), 195-201 (2001)  
MEDLINE 21601119  
REFERENCE 2 (bases 1 to 1585)  
AUTHORS Suga, H.  
TITLE Direct Submission  
JOURNAL Submitted (03-OCT-2000) Hiroshi Suga, Dept. Biophys., Grad. Sch. Sci., Kyoto Univ.; Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan (E-mail: suga@biophys.kyoto-u.ac.jp, Tel:81-75-753-4224, Fax:81-75-753-4223)  
FEATURES  
source location/Qualifiers  
1..1585  
/organism="Ephydatia fluviatilis"  
/db\_xref="taxon:31330"  
1..1585  
/gene="EFPTK150"  
8..1420  
/gene="EFPTK150"  
/codon\_start=1  
/product="protein tyrosine kinase"  
/protein\_id="BA82423.1"  
/db\_xref="GI:18146652"  
/translation="MGSCSSSPKRNNTYKTELKDKGVQVEHYPIYCKDYDSRT  
DDDLGFKGDLMIYITSEEDYDMFARLAKDTRKGYIPSTVAEYKSLDEAVFLGQIK  
RIABERLNPVAKVSEFLIRDSSESPGDSLVKQDVRHVRHRLDGLFVNG  
VWFOSLDILVHYRTQKQDGLCNLCPCLQTEVPQTAGLSRQANKMEIEDKTOIQLT  
KLQAGMLGEVMEGWNNGTTSVAVKTRPGVSVEVFLQARIVKLRHRLIOLVAVC  
TKKEPIYVMELMKYGSLEYELKEGRTLKEIYLDVIAOVASGMSLYRQONYIHCDL  
AARNILVEGHGCMKVADIGLTKYVIDONMEAGCAKRPKWTAPALIVRPTKSDVW  
SFGVLIYEITIGRPVYPIGTTHDEVLEKIQGYRMCPNCRKREYIDIMDCMHEDPA  
SRPTFTLQMLEPFRNSG"

BASE COUNT 441 a 322 c 418 g 404 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.204 Length: 1585  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x AB049595 (1-1585)

OY 1003 SerAspValTyrSerPheGlyValValLeuTYrGlu 1014  
DB 1172 TCGGATGTATGTCATTTGGGTTGCTCTTATGAG 1207

RESULT 91  
MM018084  
LOCUS MMU018084 1602 bp mRNA linear ROD 02-FEB-1996  
DEFINITION Mus musculus receptor tyrosine kinase mRNA, partial cds.  
ACCESSION U18084  
VERSION U18084.1 GI:1173482  
KEYWORDS mouse.  
SOURCE Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1602) Lickliter, J.D., Smith, F.M., Olsson, J.E., Mackwell, K.L. and Boyd, A.W.

**TITLE** Embryonic stem cells express multiple Eph-subfamily receptor tyrosine kinases

**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 93 (1), 145-150 (1996)

**MEDLINE** 96133894

**PUBMED** 8552593

**REFERENCE** 2 (bases 1 to 1204)

**AUTHORS** Lickliter, J.

**TITLE** Direct Submission

**JOURNAL** Submitted (04-DEC-1994) Jason Lickliter, Lions Laboratory, Walter and Eliza Hall Institute of Medical Research, Royal Parade, Parkville, Victoria, 3050, Australia

**COMMENT** On Feb 2, 1996 this sequence version replaced gi:1172083.

**FEATURES** Location/Qualifiers

**Source** 1. .1602

**organism** "Mus musculus"

**strain** "129/SV"

**db\_xref** "taxon:10090"

**clone** "35C15"

**cell\_line** "W9.5"

**dev\_stage** "embryonic stem cells"

**cds** <1..51602

**codon\_start** -2

**product** "receptor tyrosine kinase"

**protein\_id** "AAC52384.1"

**db\_xref** "GI:1173483"

**translation** "CPTGFYRVDMNTLRCLKCPQHSIAESGSLTCTCENGHYRANGE GPVACHTRPSPASGLTSTVQVGLSEYANTFVKNRYSGLDSSPSSAS INMHASLSGLSLKLVKRPRLQELFMASRPNGNISTELAHVINDOEHWL EPRVLRRLQPDITYIVRRTLRGLGPGSPDHEFTSPVRSLSIGRTIVATGL LIGTALIGIVFERSRGQRORQORERTNGREDKLMKLPVDAIATEDPAQCAL DFAOELDPAMLIYDTVIGEGYEGYRGALRLPSQDCTVAIKLRDTSDEGYWNNL REATINGOFNPHILRLLEGVYTKRPMIITEFMENGALDAFLKREGOLAPOLVAM LIGTASGMNCSGNYVHRDLAARNIILNCKVSDLGITRLDDFDGTYETOGCK IPIRWTPAGALAHRIETTASDVMSFG"

**BASE COUNT** 391 a 442 c 439 g 330 t

**ORIGIN**

**Alignment Scores:**

**Pred. No.:** 0.206 **Length:** 1602

**Score:** 12.00 **Matches:** 12

**Percent Similarity:** 100.00% **Conservative:** 0

**Best Local Similarity:** 100.00% **Mismatches:** 0

**Query Match:** 1.09% **Indels:** 0

**DB:** 10 **Gaps:** 0

**US-09-397-967-16 (1-1099) x MMU18084 (1-1602)**

**Qy** 941 valhisargaspleuallalargasnlleluval 952

**Db** 1400 gtccatagagacctgctccagagacatcttggtg 1435

**RESULT 92**

**HUMERKIP** 1618 bp mRNA linear PRI 29-MAY-2002

**LOCUS** Homo sapiens mRNA for large erk/cek5 tyrosine kinase, partial cds.

**DEFINITION** D37827

**ACCESSION** D37827.1 GI:1060894

**VERSION** large erk/cek5 tyrosine kinase.

**KEYWORDS** Homo sapiens 4-month fetal brain cDNA to mRNA, clone\_lib:lambda

**SOURCE** gt10.

**ORGANISM** Homo sapiens

**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**AUTHORS** 1 Chan, J. and Watt, V.M.

**TITLE** eek and erk, new members of the eph subclass of receptor protein-tyrosine kinases

**JOURNAL** Oncogene 6 (6), 1057-1061 (1991)

**MEDLINE** 91296384

**REFERENCE** 2

**AUTHORS** Iwase, T., Tanaka, M., Suzuki, M., Naifo, Y., Sugimura, H. and Kino, I.

**TITLE** Identification of protein-tyrosine kinase genes preferentially expressed in embryo stomach and gastric cancer

**JOURNAL** Biochem. Biophys. Res. Commun. 194 (2), 698-705 (1993)

**MEDLINE** 93343925

**REFERENCE** 3 (bases 1 to 1618)

**AUTHORS** Saito, T., Naohiko, S., Kitahara, M., Murata, M., Yamamoto, Y., Hori, T. and Matsuda, Y.

**TITLE** Identification of human erk gene as a putative receptor tyrosine kinase and its chromosomal localization to 1p36.1: a comparative mapping of human, mouse and rat chromosomes

**JOURNAL** unpublished

**REFERENCE** 4 (bases 1 to 1618)

**AUTHORS** Saito, T.

**TITLE** Direct Submission

**JOURNAL** Submitted (01-AUG-1994) Toshiyuki Saito, National Institute of Radiological Sciences; 9-1 Anagawa 4-chome, Inage-ku, Chiba 263, Japan (Tel:043-251-2111(ex.361), Fax:043-256-8301)

**COMMENT** On Nov 13, 1995 this sequence version replaced gi:1529066.

**Submitted** (01-Aug-1994) to DDBJ by:

Toshiyuki Saito

National Institute of Radiological Sciences

Division of Genetics

9-1 Anagawa 4-chome, Inage-ku

Chiba 263

Japan

Phone: 043-251-2111 x333

Fax: 043-251-9818.

**FEATURES** Location/Qualifiers

**Source** 1. .1618

**organism** "Homo sapiens"

**db\_xref** "taxon:9606"

**chromosome** "1"

**map** "1p36.1"

**tissue\_type** "brain"

**clone\_lib** "lambda gt10"

**dev\_stage** "4-month fetal"

**gene** 1. .1618

**gene** "ERK"

**cds** <1..1437

**gene** "ERK"

**codon\_start** =1

**product** "large erk/cek5 tyrosine kinase"

**protein\_id** "BA07073.1"

**db\_xref** "GI:1060895"

**translation** "RARTVAGYGRYSGRNWFOTMEAEYQTSIQEKLPIITSSAAGL VFLIAYVIAIVCNRRGFERADSEYTDKLOHTYSGMTGKTIYDPTVEDPNEAVR EFAKEIDISCVKIEQVIGAGEFEGVCSGLKLPKREIFVALIKTKSGQTEYQRDDEL SEASINGQFDPHPVVIHLEGVYTKSTFVMIITEFMENGSLDSEFLRQNDGQFTYQVGM LRGTAAQMKTLADMYVHRDLAARNIIVNSNLVCKVSDGSLRFLDDPTNDQVIAIEOD GGRKIPIRWTPAEALIQYRKFTSASDVMSYGYVMKEVMSYGERPYWMDTNDQVIAIEOD YRLRPMDCCPALHQLMLDCWQDRNHRPFQIVITLDMKIRNPSLKAMPPLSGI NLPLDRTIPDYTSFNTVDEMLKAIKMGQYKSEFANAGFTSDVVSQMMEDILRVGV TLGHQKKILINSIQVMRAQMNQIQSVEV"

**411**

**gene** "ERK"

**cds** /citation=[2]

**replace** "a"

**487..489**

**gene** "ERK"

**cds** /citation=[2]

**replace** "cgc"

**562**

**gene** "ERK"

**cds** /citation=[2]

**replace** "g"

**1570**

**gene** "ERK"

**cds** /citation=[2]

**replace** "g"

**unsure**

**unsure**

**unsure**

BASE COUNT 376 a 494 c 444 g 304 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.207 Length: 1618  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-397-967-16 (1-1099) x HUMERKIP (1-1618)  
 OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 |||||  
 Db 703 GTTCACCGTGACCTGGCTGCCCGAACATCTCTGTC 738  
 RESULT 93  
 LOCUS AB049594 1619 bp mRNA linear INV 27-DEC-2001  
 DEFINITION Ephradia fluvialtilis EFPTK45 mRNA for protein tyrosine kinase,  
 complete cds.  
 ACCESSION AB049594  
 VERSION AB049594.1 GI:18146649  
 KEYWORDS  
 SOURCE Ephradia fluvialtilis cDNA to mRNA.  
 ORGANISM Ephradia fluvialtilis  
 Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;  
 Haptosclerida; Spongillidae; Ephradia.  
 REFERENCE  
 AUTHORS 1 Suga, H., Katoh, K. and Miyata, T.  
 TITLE Sponge homologs of vertebrate protein tyrosine kinases and frequent  
 domain shufflings in the early evolution of animals before the  
 parazoan-eumetazoan split  
 JOURNAL Gene 280 (1-2), 195-201 (2001)  
 MEDLINE 21601119  
 REFERENCE 2 (bases 1 to 1619)  
 AUTHORS Suga, H.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-OCT-2000) Hiroshi Suga, Dept. Biophys., Grad. Sch.  
 Sci., Kyoto Univ., Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto, Kyoto  
 606-8502, Japan (E-mail: suga@biophys.kyoto-u.ac.jp,  
 Tel:81-75-753-4224, Fax:81-75-753-4223)  
 FEATURES  
 source location/Qualifiers  
 1..1619  
 /organism="Ephradia fluvialtilis"  
 /db\_xref="taxon:31330"  
 1..1619  
 /gene="EFPTK45"  
 112..1599  
 /gene="EFPTK45"  
 /codon\_start=1  
 /product="protein tyrosine kinase"  
 /protein\_id="BAB82422.1"  
 /db\_xref="GI:18146650"  
 /translation="MGSLSSPDSSDKSKGVEITNKGKROPEPRKARQAPRVQZ  
 VKPYVYGVYDSTDDLSFKKQDLRIITSTBDGDMFRASKOTGEGITPSVYA  
 EKSIDAEMLFKIKRYAEKILMOSNQISFLIRSETPGDFSLVSDODRVRA  
 YRIRLEDGSLFVTRSTFGHLEHVEHYKNTDOLCCNLVPCLOAEKPTAGLSRQ  
 ANEWEIKKOIKIKTKLAGOFGFGEVWNGTTSVAVKTLKPGTMSVEFLDASI  
 MKKLHPKLIOLVACTKEEPIYIVTELKYGSLLEYLRDGRITLKIDRLADIAOVA  
 SGMSTLEQOANYIHDLARNTLVGSHGICVADGGLAVDEIEITYEANTGAKPKIKNT  
 APEAMVNRFTIKSDVMSFGVLYEITTYGRRPPIFGMTNPVLEKIQOQYRMPCPANC  
 PKQYHDVLDGMRDPPASRPTEFLQWLEEFENSEGRDPAHL"  
 BASE COUNT 449 a 363 c 432 g 375 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.208 Length: 1619  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 9 Gaps: 0

DB: 3 Gaps: 0  
 US-09-397-967-16 (1-1099) x AB049594 (1-1619)  
 OY 1003 SerAspValTrpSerPheGlyValIleuTyGlu 1014  
 |||||  
 Db 1327 TCAGATGTGTGTCGTCGTTGGGCTTGCTTTATGAG 1362  
 RESULT 94  
 LOCUS HSM801500 1619 bp mRNA linear PRI 18-FEB-2000  
 DEFINITION Homo sapiens mRNA: cDNA DKFZp434C1418 (from clone DKFZp434C1418);  
 partial cds.  
 ACCESSION AL133666  
 VERSION AL133666.1 GI:6599298  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS 1 (bases 1 to 1619)  
 Othenwaelder, B., Oberwälder, B., Mewes, H.W., Gassenhuber, J. and  
 Wiemann, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-DEC-1999) MIPS, Am Klopfersplitz 18a, D-82152  
 Martinsried, GERMANY  
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by Medigenomix (Martinsried/Germany) within the cDNA  
 sequencing consortium of the German Genome Project. This clone  
 (DKFZp434C1418) is available at the RZPD in Berlin. Please contact  
 the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
 information about the clone and the sequencing project is available  
 at <http://www.mips.biochem.mpg.de/proj/cDNA/>.  
 FEATURES  
 source location/Qualifiers  
 1..1619  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp434C1418"  
 /tissue\_type="testis"  
 /clone\_lib="434 (synonym: htes3). Vector pSPORT1; host  
 DH10B; sites NotI + SalI"  
 /dev\_stage="adult"  
 343..1347  
 /gene="DKFZp434C1418"  
 343..1347  
 /note="Strong similarity to receptor-like tyrosine kinase  
 Etk-2"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CA863775.1"  
 /db\_xref="GI:6599298"  
 /translation="MKDSPQVTRKLYLNKWFIFASASDMAEQOILVATAVAGS  
 FTLLVITLLEFLILGRQWYIKAKMKEERKRNHQLRFPKIKYIPDVEDPS  
 LAHVEFAKEIDPSRIETRIYVIGAEFGEVSGSRKTEIPVATIKTLGGHMDROR  
 RDLFRKASIKGQFDHPNIRILEGVTRSPPAIGVAFKPSPLRAGFLNIOAPHPYP  
 GGSLPRIPRIPAGRPVMTVVMEMSGISDPLRRHDSHFYIQLVGMGRGLASGSKTYS  
 DMGVVHMDLARNITLVNSLVCKVSDGLSRVLEDDPEAAVTTTDLFQTLNLNCTSA  
 " polyA\_signal  
 BASE COUNT 508 a 338 c 362 g 411 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.208 Length: 1619  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x HSM801500 (1-1619)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 1183 GTTCATCGAGACCTAGCGCTCGGAATATCTGCTC 1218

## RESULT 95

MMSEK3

LOCUS M.musculus Sek-3 mRNA. 1635 bp mRNA linear ROD 27-APR-1995

ACCESSION X76011.1 GI:460055

KEYWORDS Sek gene.

SOURCE Mus musculus.

ORGANISM

REFERENCE 1 (bases 1 to 1635)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Becker, N., Seltanidou, T., Murphy, P., Mattei, M.G., Topilko, P., Nieto, M.A., Wilkinson, D.G., Charnay, P. and Giliardi-Hebenstreit, P.

TITLE Several receptor tyrosine kinase genes of the Eph family are segmentally expressed in the developing hindbrain

JOURNAL Mech. Dev. 47 (1), 3-17 (1994)

MEDLINE 95034306

PUBMED 7947319

REFERENCE 2 (bases 1 to 1635)  
Becker, N.  
Direct Submission  
Submitted (11-OCT-1993) N. Becker, Unite Inserm U368 Ecole Normale, Superieure, Laboratoire de Biologie Moleculaire, du Developpement-Patrick Charnay, U368 INSERM 46 Rue d'Ulm 75005 Paris, FRANCE

AUTHORS

TITLE Location/Qualifiers

JOURNAL 1. 1635

MEDLINE /organism="Mus musculus"

PUBMED /strain="BALB/c"

REFERENCE /db\_xref="taxon:10090"

AUTHORS /chromosome="4"

JOURNAL /map="D.E"

FEATURES /clone="15"

SOURCE /tissue\_type="brain"

1. 1443

/gene="Sek-3"

CDS &lt;1..1443

/gene="Sek-3"

/protein\_id="CAA53598.1"

/db\_xref="GI:460056"

/db\_xref="MGI:96611"

/translation="VARTVAGYGRYSGMYFQMTTEAYQTSIREKLLPLIVSSAAG  
LVFLIAVVIVICNRGRFPERADSEYDQLQHTSGHMTGKMYIDPFTYEDPNEA  
VREFAKEDISCVCLEQVIGAGEGEVSGHLKLPGRKEIFVAIKTKSGTYEKORD  
FLSEASIMGQDHPVHILBGVVMKSTPVMIITEPMENGSDFLRONDGGQFVIOY  
GMLKGIAGMKYTLADMTVYHDLARNLIVSNLYCKKSDGLSRELPDPTSDPTYS  
ALGKGIPIRMVPEAIQYRFTSADVSTGIVMEVSYSEPRVMDTNDVYNAIE  
QDYRLPMPDPSALHQLMDQCKDRHMRKFGQIVNTDKMLRNPSLKAAPLSS  
GINLPIDRLTIPDYSEFNTVDEMLEAIKMGQYKESFANAGFTSPDVVSQMMEDILRV  
GVTLAGHOKITLINSIQVRAQMNQIQSYEV"

misc-feature 1..99

misc-feature /gene="sek-3"

misc-feature /note="extra cellular region"

misc-feature /gene="Sek-3"

misc-feature /note="trans membrane region"

misc-feature 178..1443

misc-feature /gene="sek-3"

Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.09% Indels: 0

DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x MMSEK3 (1-1635)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 709 GTGCACCGTAGACTTGTCTGCTCGAACAATCTCTGCTC 744

RESULT 96

LOCUS MMU05210 1651 bp mRNA linear ROD 08-JUN-1994

ACCESSION U05210

KEYWORDS Mus musculus protein tyrosine kinase Ctk mRNA, complete cds.

SOURCE house mouse.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1651)  
Klages, S., Adam, D., Class, K., Fargnoli, J., Bolen, J.B. and Penhallow, R.C.

AUTHORS Ctk: a protein-tyrosine kinase related to Csk that defines an enzyme family

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (7), 2597-2601 (1994)

MEDLINE 94195789

PUBMED 7511815

REFERENCE 2 (bases 1 to 1651)  
Penhallow, R.C.  
Direct Submission  
Submitted (14-JAN-1994) Robert C. Penhallow, Molecular Biology, Bristol-Myers Squibb Pharmaceutical Research Institute, Route 206 and Provincetown Road, Princeton, NJ 08543-4000, USA

AUTHORS Location/Qualifiers

JOURNAL 1. 1651

MEDLINE /organism="Mus musculus"

PUBMED /db\_xref="taxon:10090"

REFERENCE /tissue\_type="brain"

AUTHORS /clone\_lib="murine brain, lambda ZAP"

JOURNAL /dev\_stage="adult"

FEATURES /function="protein tyrosine kinase"

SOURCE /codon\_start=1

117..1514

/product="ctk"

/protein\_id="AA18829.1"

/db\_xref="GI:450233"

/translation="MPTRMAGTCYQCMKCNRPKGEIAFRKGDWVTLFACEDKSM  
YRANHSSGQGLIAAALRHGEALSTDPKLSMPFHKISQGEALQOLQPEDGFL  
LVRSARHPGGYVLCVSGFRVHYRLHNDGHLTIDEAVCFMLKMDVHYTKDKKA  
ICTKLVPRRKGQAKSAEEELAKAGMLDLDLQHLTGALQIEGEVAVLQGEVIGQKVA  
VKNKICDVTADAFIDETAVMTKLOHRLNVLGLVILHNGLYIWEHVSQNLVNFLLT  
RGRALVSTSQLQFALHVAEGMELESEKTIYHRLDARNLIVSBDLVAKYSDGLAKA  
ERKLDSSRLPRKMTAPALKNRFESEKSDVMSGVILWVESYGRAPYKMSLKEYS  
EAVKGRMEPPDCCPSVHTLMGSCHEABEAPRRPPRIYKIKRGLRISGVASAPAG  
GQEAEGSAFPTSDDP"

misc-feature 365 a . 445 c 530 g 311 t

misc-feature /gene="sek-3"

misc-feature /note="extra cellular region"

misc-feature /gene="Sek-3"

misc-feature /note="trans membrane region"

misc-feature 100..177

Alignment Scores: 0.211 Length: 1651

Pred. No.: 12.00 Matches: 12

Score: 100.00% Conservative: 0

Percent Similarity: 100.00% Mismatches: 0

Best Local Similarity: 100.00% Indels: 0

Query Match: 1.09% Gaps: 0

DB: 10

US-09-397-967-16 (1-1099) x MMU05210 (1-1651)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 941 GTTCATCGAGACCTAGCGCTCGGAATATCTGCTC 1218



```

Db 1035 GTGCACCGGACCTGGCTGCTCGAATCTCTGTC 1070
RESULT 97
LOCUS S77473
DEFINITION 1656 bp mRNA linear ROD 25-AUG-1995
ACCESSION S77473
VERSION S77473
KEYWORDS S77473.1 GI:957295
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Ersler, M.A., Krivtsov, A.V., Krockova, A.V., Bellavskii, A.V. and
Visser, J.V., 1995.
TITLE VNK-a new gene for nonreceptor protein-tyrosine kinase, expressed
JOURNAL in the murine brain and hematopoietic system
MEDLINE Doml. Akad. Nauk. 339 (5), 679-683 (1994)
PUBMED 95178994
REMARK 7874011
Genbank staff at the National Library of Medicine created this
entry (NCBI gi95716512) from the original journal article.
COMMENT This sequence comes from Fig. 1.
FEATURES
SOURCE A base 'a' has been added at position 1408 to accommodate the amino
acid sequence given.
Location/Qualifiers
1..1656
/organism="Mus sp."
/db_xref="taxon:10095"
1..1656
/gene="VNK"
2..1537
/note="nonreceptor protein-tyrosine kinase"
2..1537
/gene="VNK"
/note="nonreceptor protein-tyrosine kinase: this sequence
comes from Fig. 1; Protein sequence is in conflict with
the conceptual translation: mismatch(421[E->Q])"
/codon_start=1
/db_xref="GI:957296"
/protein_id="AAB33566.1"
/translation="MRSIASCSELSHWVLDPLAGTLRSGPRKVSRLFGANHPAPAA
ARPTQRMABPTGCTCKCNSRPRKGEALFRKGDWTTIEACEDSWRAKHGSGOE
GLTAAALRORELSIDPKSLMPHFGKISQCEAIQQLQPEDLFLVRESARHGD
YLVCSFEGRDVHYRVLHNDHGLTIDEAVCFNLDVNEHYTRKGAICTKLVKPRK
OGAKSABEELAKAGWLDLQHLTLGAQIGEGEGLVGEVGLGKVAANKICDVTQA
ARLDETVATKLOHRNLVRLGLVLIHGLYIWEHVSNGNLVNFRLTRGALVSTQA
LQDALHVAAGMEYLESKLVHRDLAARNILVEDLVAKSDEGLAKARKGIDSSRLP
VKWTAEALKNRFSKSDVSVFVGLVMEVFYGRAPYPRKMSLKEVSRVAVKGYMPE
PDGCPGSHVTLMSGSCWEAPRRPFRKIVKELGRDVSVCSAAGGDEAGSAFTRS
ODP"
BASE COUNT 355 a 447 c 536 g 317 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 0.211 Length: 1656
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0
US-09-397-967-16 (1-1099) x S77473 (1-1656)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 1061 GTGCACCGGACCTGGCTGCTCGAATCTCTGTC 1096
RESULT 98
LOCUS BC008655
1658 bp mRNA linear PRI 12-JUL-2001

```

```

DEFINITION Homo sapiens, clone IMAGE:3852708, mRNA, partial cds.
ACCESSION BC008655
VERSION BC008655.1 GI:14250439
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1658)
TITLE Strausberg, R.
JOURNAL Direct Submission
Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NHI-MGC Project URL: http://mgc.ncl.nih.gov
REMARK Contact: MGC help desk
COMMENT Email: cgapds-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: Villalón, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAX Plate: 13 Row: a Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
FEATURES
SOURCE Location/Qualifiers
1..1658
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3852708"
/tissue_type="Colon, adenocarcinoma"
/clone_id="NHI_MGC_65"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
<1..763
/codon_start=2
/product="Unknown (protein for IMAGE:3852708)"
/protein_id="AAH08655.1"
/db_xref="GI:14250440"
/translation="AAGMKYLANMYVHRDLAARNILVNSNLCKYSDGSLRVLEDD
PEATYTSGGKIPITRMVAPPAISYRKFTSADVWSGIYMWEMVTIGERPVLSHE
VKAINDGFRLLPMDPCSAIYIOLMQOQDEARRPRADIVSLIDKIRAPDSIKT
LADPDPRVSRILRSTSGSEVPRTVSEWLESTIKMOQYEHFMAGYTAIEKVVOMTN
DDIKRGVIRLPGHOKRIYAVSLGLKDVNTVGIPI"
BASE COUNT 350 a 488 c 443 g 377 t
ORIGIN
Alignment Scores:
Pred. No.: 0.212 Length: 1658
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0
US-09-397-967-16 (1-1099) x BC008655 (1-1658)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 38 GTGCACCGGACCTGGCTGCTCGAATCTCTGTC 73
RESULT 99

```

MUSNTK  
LOCUS 1734 bp mRNA linear ROD 01-JUL-1994  
DEFINITION Mus musculus tyrosine protein kinase (Ntk) mRNA, complete cds.  
ACCESSION L27738  
VERSION L27738.1 GI:507289  
KEYWORDS tyrosine protein kinase.  
SOURCE Mus musculus (strain BALB/c, sub-species domesticus) day 16 fetus thymus cDNA to mRNA.

ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1734)  
Chow, L.M., Jarvis, C., Hu, Q., Nye, S.H., Gervais, F.G., Veilleux, A. and Matis, L.A.  
Ntk: a Csk-related protein-tyrosine kinase expressed in brain and T lymphocytes  
Proc. Natl. Acad. Sci. U.S.A. 91 (11), 4975-4979 (1994)

JOURNAL  
MEDLINE  
PUBMED  
94255451  
8197166

FEATURES  
source  
1..1734  
/organism="Mus musculus"  
/strain="BALB/c"  
/sub\_species="domesticus"  
/db\_xref="taxon:10090"  
/tissue\_type="thymus"  
/dev\_stage="day 16 fetus"  
/germline  
1..1734  
/gene="Ntk"  
1..97  
/note="Ntk"  
/note="putative"  
98..1615  
/gene="Ntk"  
/note="amino acid feature: SH2 domain, bp 455..787; amino acid feature: SH3 domain, bp 248..454; amino acid feature: carboxy-terminal tail, bp 1616..1756; amino acid feature: tyrosine kinase domain, bp 788..1525; amino acid feature: unique domain, bp 98..247"  
/codon\_start=1  
/evidence="experimental"  
/product="tyrosine protein kinase"  
/protein\_id="AAB59677.1"  
/db\_xref="GI:507290"  
/translation="MARSSRSVSMIAFEQESRDLPVPSPLFGAMHPAPAAAMPTRWAPGTOMCTKCEKSPKPELAFKRGDMYTIIEACEDKMYRAKHHSQOEGLIAAA LRGREALSTDPKLSLMPWFHGIKSGOEAIQOOLPPEDGFLVRESARRHGDVLCVSE GRDVIHYRVLHRRDGLTIDEAVCFNLMDVETIKDKGICTKLYKPRKRGKASAE EELAKAGMLIDGLTGAOIEGEGAVLQGYLGQKAVANIKODVIAQAFLEBETA VETRLQHRNLVRLGLVILHGLYIVMEHVSQKMLVNEFLTRGRALVSTQLQFLAHV ABGEVYTESKIKVHRDLAARNLIVSEDLVAKVDFGLAEKRGIDSSLPYKWAPE ALKNGRSSKSDVMSFCVILMEVEFSYGRAPYKMSLKEVSEAVEKGYRMEPPDGGGS VHLNMGSCWEAEFAPRRPPRKIVYKELGRLEIRLGSVAPAGCGAEEASATRSQDP" 226-227

misc\_difference  
226-227  
/gene="Ntk"  
/standard\_name="3 nucleotide insert"  
/note="approximately 15% of cellular mRNAs contain this insert"

3'UTR  
1616..1734  
/gene="Ntk"  
1709..1713  
/gene="Ntk"  
1734  
/gene="Ntk"

polyA\_site  
1734  
/gene="Ntk"

BASE COUNT 362 a 480 c 557 g 335 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.219 Length: 1734  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x MUSNTK (1-1734)  
QY 941 ValHISArgAspLeuAlaAlaArgAsnIleLeuVal 952  
|||||  
Db 1136 GTGCACCGGAGCTGCTGCTCGACATCTCGTGC 1171

RESULT 100  
S71669 1744 bp mRNA linear PRI 24-OCT-1994  
LOCUS lsk-leukocyte carboxyl-terminal src kinase related gene [human,  
DEFINITION peripheral T cells, mRNA, 1744 nt].  
S71669  
S71669.1 GI:559593  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens peripheral T cells.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
1 (bases 1 to 1744)  
McVicar, D.W., Lal, B.K., Lloyd, A., Kawamura, M., Chen, Y.Q., Zhang, X., Staples, J.E., Ortaldo, J.R. and O'Shea, J.J.  
Molecular cloning of lsk, a carboxyl-terminal src kinase (csk) related gene, expressed in leukocytes  
Oncogene 9 (7), 2037-2044 (1994)

JOURNAL  
MEDLINE  
PUBMED  
94268844  
7516063

REMARK  
GenBank staff at the National Library of Medicine created this entry [NCBI 91bbsq 151068] from the original journal article.  
This sequence comes from Fig. 1.

FEATURES  
source  
1..1744  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..1744  
/gene="lsk"  
/note="leukocyte carboxyl-terminal src kinase related gene"  
125..1522  
/gene="lsk"  
/note="leukocyte carboxyl-terminal src kinase related gene: This sequence comes from Fig. 1"  
/codon\_start=1  
/protein\_id="AAB30995.1"  
/db\_xref="GI:559594"  
/translation="MPTRKMAPGTGCTTCCKEHTRRPKRGELAFKRGVNTILCEKNS WYRVKHNHSGQEGILAAAGALREKALADADRLSLMPWFHGIKSGOEAVQLOPPEDGL FLVRESGHPGDVLCVTFGRDVIHYRVLHRRDGLTIDEAVCFNLMDVETIKDKGICTKLYKPRKRGKASAE EELAKAGMLIDGLTGAOIEGEGAVLQGYLGQKAVANIKODVIAQAFLEBETA VETRLQHRNLVRLGLVILHGLYIVMEHVSQKMLVNEFLTRGRALVSTQLQFLAHV ABGEVYTESKIKVHRDLAARNLIVSEDLVAKVDFGLAEKRGIDSSLPYKWAPE ALKNGRSSKSDVMSFCVILMEVEFSYGRAPYKMSLKEVSEAVEKGYRMEPPDGGGS VHLNMGSCWEAEFAPRRPPRKIVYKELGRLEIRLGSVAPAGCGAEEASATRSQDP" 226-227

BASE COUNT 354 a 534 c 582 g 274 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.22 Length: 1744  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x S71669 (1-1744)  
QY 941 ValHISArgAspLeuAlaAlaArgAsnIleLeuVal 952  
|||||  
Db 1043 GTGCACCGGAGCTGCTGCTCGACATCTCGTGC 1078

Tue Apr 29 06:04:25 2003

Search completed: April 28, 2003, 21:08:33  
Job time : 7366 secs

us-09-397-967-16.oligo.rge

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 28, 2003, 13:46:19 ; Search time 6835.27 Seconds

(without alignments)  
4679.258 Million cell updates/sec

Title: US-09-397-967-16  
Perfect score: 5860  
Sequence: 1 MAPSEEPFLIPQRSCSLSS.....RPAFATLSPQLDPLMRGRPG 1099

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O/cgnt2.1/USPTO.spool/US09397967/rnat.26042003.182312.10714/app.query.fasta.1.1486  
-DB=GenEmbl -OPM=fastp -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.coi -LIST=45  
-DOCALIGN=200 -THR=SCORE=pcr -THR=MAX=100 -THR=MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcr -NOR=ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=2000000000  
-USBR=US09397967\_@CGN\_1\_15398\_@rnat.26042003.182312.10714 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: GenEmbl: \*  
1: gb\_da: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_ov: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_da: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5854	99.9	3723	10 MUSPYKIN	L32955 Mouse prote
2	5687.5	97.1	3528	10 MUSJAK3A	L40172 Mus musculu
3	5380.5	91.8	3778	10 RATJAK3	D28508 Rat mRNA fo
4	5223	89.1	4016	10 MUSJAK3H	L33768 Mus musculu
5	4684	79.9	4064	6 AX203081	AX203081 Sequence
6	4604	79.9	4064	9 HSU09607	U09607 Human JAK f
7	4597	78.4	3620	9 HSU31601	U31601 Human tyros
8	4091.5	69.8	218208	2 AC073700	AC073700 Mus muscu
9	4091.5	69.8	223734	2 AC073750	AC073750 Mus muscu
10	3935	67.2	10326	10 NMU71201	U71201 Mus musculu
11	3336	58.9	3644	5 AF034576	AF034576 Gallus ga
12	3228.5	55.1	2914	6 AX203084	AX203084 Sequence
13	2713	46.3	13562	9 HSU70065	U70065 Human JAK3
14	2637.5	45.0	4812	4 AB006011	AB006011 Sus scrof
15	2636.5	44.9	3435	6 RNU13396	U13396 Rattus norv
16	2629	44.9	3435	6 AF058925	AF058925 Sequence
17	2629	44.9	4161	9 AF005216	AF005216 Homo sapi
18	2628	44.8	3500	9 AF001362	AF001362 Homo sapi
19	2621	44.7	5117	6 AX354568	AX354568 Sequence
20	2621	44.7	5117	6 AF058925	AF058925 Homo sapi
21	2619.5	44.7	3629	6 AR136148	AR136148 Sequence
22	2619.5	44.7	3629	6 AR143945	AR143945 Sequence
23	2619.5	44.7	3629	6 I92575	I92575 Sequence 8
24	2613.5	44.6	3629	10 MUSPYK	L16956 Mouse prote
25	2606.5	44.5	4387	5 DRE5690	AJ005690 Danio rer
26	2523.5	43.1	3739	5 AF148993	AF148993 Cyprinus
27	2511	42.8	4042	4 AB036337	AB036337 Sus scrof
28	2460.5	42.0	2881	9 BC028068	BC028068 Homo sapi
29	2382.5	40.7	41006	9 AC007201	AC007201 Homo sapi
30	2360	40.3	3495	6 AR048200	AR048200 Sequence
31	2360	40.3	3495	6 AR068116	AR068116 Sequence
32	2360	40.3	3495	6 AR071176	AR071176 Sequence
33	2360	40.3	3495	6 AR177639	AR177639 Sequence
34	2360	40.3	3495	6 I62349	I62349 Sequence 2
35	2360	40.3	3495	6 I62349	I62349 Sequence 2
36	2350	40.1	21323	4 AB036336	AB036336 Sus scrof
37	2239	39.9	3720	5 AF513860	AF513860 Homo sapi
38	1967.5	33.6	5065	4 AB036335	AB036335 Sus scrof
39	1915.5	32.7	4191	10 S63728	S63728 Mus sp. JAK
40	1908.5	32.6	4191	10 S63728	S63728 Mus sp. JAK
41	1904	32.5	3429	6 AR136149	AR136149 Sequence
42	1904	32.5	3429	6 AR143946	AR143946 Sequence
43	1904	32.5	3429	6 I92576	I92576 Sequence 10
44	1904	32.5	3541	6 AX329654	AX329654 Sequence
45	1904	32.5	3541	6 AX409732	AX409732 Sequence

RESULT 1

ALIGNMENTS

MUSPTRYKIN 3723 bp mRNA linear ROD 21-SEP-1994  
LOCUS  
DEFINITION Mouse protein tyrosine kinase (jak3) mRNA, complete cds.  
ACCESSION L32955  
VERSION L32955.1 GI:529238  
KEYWORDS interleukin 2; interleukin 4; Janus kinase; protein tyrosine kinase.  
SOURCE Mus musculus (strain BALB/c, sub.species domesticus) cDNA to mRNA.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 3723)  
AUTHORS Mithun,B.A., Silvenoinen,O., Miura,O., Lai,K.S., Cwik,C., Liu,E.T. and Inhe,N.  
TITLE Involvement of the Jak-3 Janus kinase in signalling by interleukins 2 and 4 in lymphoid and myeloid cells  
JOURNAL Nature 370 (6485), 153-157 (1994)  
MEDLINE 94294024  
PubMed 8022486  
COMMENT This entry has been reported under the accession number L32955 in the Nature article.  
FEATURES  
source location/Qualifiers  
1..3723  
/organism="Mus musculus"  
/strain="BALB/c"  
/sub\_species="domesticus"  
/db\_xref="taxon:10090"  
/cell\_line="Abelson transformed pre-B-cell line 22D6"  
/cell\_type="lymphoid and myeloid"  
1..3723  
/gene="jak3"  
1..12  
/gene="jak3"  
13..3312  
/gene="jak3"  
/standard\_name="Janus kinase, mjak3"  
/codon\_start=1  
/evidence=experimental  
/product="protein tyrosine kinase"  
/protein\_id="AA021565.1"  
/db\_xref="GI:529239"  
/translation="MAPPSEETPLIPORSGSLSSSEACALHYLLPPRGRPPQRLSES  
FGDYLEDLCVRAKAKALIPVHSLFALATEDSCWPSHIFCIEDVDVQVLYRL  
RFEPDMFGLTCHRFGRKDLTSLIDLHLVLEHLPADHRDLSVGRPLVGLSMKEG  
EFLVAVLDLQMAKEORQREGELKTIVSYACAPSLRDVIOCONFTVRRIRRTVY  
LALLPGRLPGRPYALMAKYTLIDLERLHPATTEFRVGLGAEPEGLRVAGDNGI  
PWSNDELFOFCDFPEPVDSINAPRVGAGRHRLVTVRMGHILEAPEGLPEA  
LSFVALVDGYFRLLTODSRHYCKEYAPRLLIEEDVCHGTTITDFAIHKAKAGSLP  
GTYILRRSPDYDSFLTACVQTPGLPYKQCLLRQDSGAPSLVGLSQPHRSLRELL  
AACWNSGLRVGALYLTSCCAPRKESNLIYVRGDNAPAPQCSFSCALQULSF  
HTPTDSLMEHNLGHGSFTKIFRGSREVVDETHDEVLKWDNRHRCMSFLE  
AASLSQVSPHYLVLLHGVCMAGDSIMVQEEVYGAIDMYLRKHLLVASMKLOVTK  
OLAYALNYLEDKGLPHGNVSAKXYLLAREGGDGNPFIKLSDPGSPVLVLEMLTDR  
IPWABECLQENOTICLEADKKGEGATTWEVFORPAITSLIEPAKKIKVEYDQGLP  
ALKWTELGLITQCMAYDPGRPSFRLRLKDLNGLITSDYELSDPTTGITSPRDEL  
VAGQIVACODPAFEERHLKYISLGGNGSGVLCRYDPLDGLVAVKQLOHS  
VPDQORDFOREIOLIKALHSDFIYKRVSGYSGQSLRVMEVYLPSCILNDLQHR  
GLHTRLLFLMAOICGMVEYLGARCVHDLAARNLIVSESHAVYIAFGAKLPLIG  
KDYVVRERGGSPIFWVAPESLSDNIFRSDSVMSGVIYELVETFCYCDKSCSPSEFL  
RMNGPERGPPICRLLELLABRRRLPPPTCTEYQELMQLCMANEHPDRAPATLSP  
QIDPLMRGPPG-  
mat\_peptide 13..3309  
/gene="jak3"  
/product="protein tyrosine kinase"  
/standard\_name="Janus kinase, mjak3"  
/evidence=experimental  
3'UTR 3313..3723  
/gene="jak3"  
3723  
/gene="jak3"  
polyA\_site  
BASE COUNT 711 a 1162 c 1079 g 771 t  
ORIGIN 7

## Alignment Scores:

Pred. No.:	0	Length:	3723
Score:	5854.00	Matches:	1098
Percent Similarity:	99.91%	Conservative:	0
Best Local Similarity:	99.91%	Mismatches:	1
Query Match:	99.90%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-16 (1-1099) x MUSPTRYKIN (1-3723)

QY	1	MetAlaProProSerGIuGIuThrProLeuILeProGlnArgSerSerLeuSerSer	20
DB	13	ATGGACCTCTCAAGTGGAGACACCTCTGATCCCTGAGCGCTTGGACCTCTCAATCC	72
QY	21	SerGIuAlaGIuAlaLeuHisValLeuLeuProProArgGIuProGIuProGIuArg	40
DB	73	TCAAGGACGAGACCTTGATGCTCTCTCCCGGGGACCTGGGCTCCACACGA	132
QY	41	LeuSerPheSerPheGIuAspTYrLeuAlaGIuAspLeuGysValArgAlaAlaGysAla	60
DB	133	TTGTCATCTCTTTGGGACTACTGCTGAGGATTTATGTGGCAGCTCCACAGGCC	192
QY	61	CysGIuAlaLeuProValTYrHisSerLeuPheAlaLeuAlaThrGIuAspPheSerCys	80
DB	193	TGTGCTATCTGCTGCTGTTATCATTCGCTTTCCTGCGCATGAGACTTCTTCC	252
QY	81	TrpPheProProSerHisIlePheCysIleGIuAspValAspThrGIuValLeuValTYr	100
DB	253	TGCTTCCCGCCACGACCATCTTCTGCATAGAGACGACGACATCAAGCTTGTCTAC	312
QY	101	ArgLeuArgPheTYrPheProAspTrpPheGIuLeuGIuThrCysHisArgPheGIuLeu	120
DB	313	AGGCTACCTTTTATTTCCCGACTGTTGGCTGGAGACATGACCGCTTGGGCTG	372
QY	121	ArgIuAspLeuThrSerAlaIleLeuAspLeuHisValLeuGIuHisLeuPheAlaGln	140
DB	373	CGCAAGATTGGACGACGACCATCTTCTGACTTACATGTTTAGAACAATCTTTGGCTAC	432
QY	141	HisArgSerAspLeuValSerGIuArgLeuProValGIuLeuSerMetIuGIuGlnArg	160
DB	433	CACGCACTGACGTGCTGAGTGGGCGCTCCCGGGGCTTACACATGAAAGACAGGGA	492
QY	161	GIuPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGIuAlaGlnArg	180
DB	493	GAGTTCGAGCGTGGCGCTGCTGAGCTTGGCCAGATGCTCTGAGCAGGCCAGCGC	552
QY	181	ProGIuGIuLeuLeuIuSerThrValSerTYrIuSAlaCysLeuProProSerLeuArgAsp	200
DB	553	CCAGAGAGCTGCTGAACAGCGTACGTTACAAGCTGTGCGGCCAGCCTCGCGAT	612
QY	201	ValIleGlnGIuGlnAspPheValThrArgArgArgIleArgArgThrValValLeuAla	220
DB	613	GTGATCCAGGGCCAGAACCTTCTGTGACGACGAGGCCATCCGAGGACCGTGTGCGG	672
QY	221	LeuLeuProCysGIuArgLeuProGIuArgProTYrAlaLeuMetAlaIuSerTYrIleLeu	240
DB	673	CTGCTGCGCTGTGCTGCGCTGCGACGCGGACGCGTGCATGCGCAAGATATCTTG	732
QY	241	AspLeuGIuArgLeuHisIleProAlaAlaThrThrGIuThrPheAlaValGIuLeuProGIu	260
DB	733	GACCTGAGACGGCTACATCCAGCGGCCACACGAGACCTTCCGTGGGGCTCCGGGCG	792
QY	261	AlaGlnGIuGIuProGIuLeuLeuArgValAlaGlyAspAspGlyIleProTrpSerSer	280
DB	793	GCCCGAGAGACCGGGGCTTCTGCTGGAGGAGGACACGCAATCTTGGAGCTCG	852
QY	281	AsnAspGIuLeuPheGlnThrPheCysAspPheProGIuIleValAspValSerIleAsn	300
DB	853	AACGACGACCTTCTCCAGACCTTGTGACTTCCGGAATGTCGAGTGCATCAAT	912
QY	301	GlnAlaProArgValGIuProAlaGIuGIuHisArgLeuValThrValThrArgMetAsp	320
DB	913	CAGGCCCGACGCTGTGGGTCCGCGAGGAGACCGGCTGTGCTACACGAGATGAC	972

QY 321 GLYHISILEUGLUAAGLUPHeproGlyLeuProGluAlaLeuSerPheValAlaLeu 340  
 DB 973 GGGCAATCTCGTAACGGGAGTTTCCGGGGCTCCGTGAAGCGCTGCTTTCGGGGCTTC 1032  
 QY 341 VALASPGLYTPheArgLeuIleCysAspSerArgHisTPheCysLysGluValAla 360  
 DB 1033 GTGGATGGGTACTTCCGGCTGATCTCGACTCCAGGCATATTATTGTGCAAGAGGTGGGG 1092  
 QY 361 PROProArgLeuLeuGluGluGluAlaAspValCysHisGlyProIleThrLeuAspPhe 380  
 DB 1093 CCGCCAAAGGTGCTGGAGGAGGAGGCGGAGGTGTCATGTGACCCATCACTAGACTTT 1152  
 QY 381 ALAILEHISLysLeuLysAlaAlaGlySerLeuProGlyThrIleThrLeuArgArgSer 400  
 DB 1153 GGCATTCACAGCTGAAGGCGCTGGCTCCCTCCAGGACCCATATTCTCCGCGCCAGC 1212  
 QY 401 PROGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420  
 DB 1213 CCGCAGGACTATGACAGCTTTCTTTACCGCGCTGCTGACAGACTCTCTTGGCCCCGAC 1272  
 QY 421 TYTLyGLYCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSer 440  
 DB 1273 TACAAAGGCTGCTCATCCGCCAGACCCAGCGGGCTTCTCCCTGGTTGGCTCAGC 1332  
 QY 441 GlnProHisArgSerLeuArgLysLeuLeuAlaAlaCysTPAsnSerGlyLeuArgVal 460  
 DB 1333 CAGCCCCACAGAAAGCTCGGGAGTGTTCAGAGCTGCTGGAAATTCTGGGCTGGCAGTA 1392  
 QY 461 ASPArgLysAlaAlaLeuTyrIleThrSerCysGlyAlaProArgProLysGluLysSerAsn 480  
 DB 1393 GAGGGGTGCGCCCTGACTTAACATCTCTGCTGCTCCAGACCCCAAGGAAAAAGTCCAAAT 1452  
 QY 481 LEUleuValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500  
 DB 1453 TTGATGCTGGTGGCAAGGGGCTGCAACCCCGGGCTGCGCTGCTCCCGTCTCTTC 1512  
 QY 501 CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTPHisGlu 520  
 DB 1513 TGTGGCTGACACAGCTGAGTCTCCACAAATTCACACGAGCAGCTGAGTGGCAGCAG 1572  
 QY 521 AsnLeuGlnHisGlySerPheThrLysIlePheArgGlySerArgArgLysValAlaAsp 540  
 DB 1573 AACCTGGGTCAGGGTCTTTTACCAAGATCTTCCTGGCAGCAGGGGGAGGTGCGGGAT 1632  
 QY 541 GLyGluThrHisAspSerGluValLeuLeuLysValIleAspSerArgHisArgAsnCys 560  
 DB 1633 GGTGAGACACATGACTCGGAAGTCCCTGGAAGTCAATGAGTCCAGACATGGAACTGC 1692  
 QY 561 MetGlnSerPheLeuGlnAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580  
 DB 1693 ATGGAGTCTTTTCTGGAACCCGCAAGCTTGATGACCAAGTATCTTACCCTGACCTGGTG 1752  
 QY 581 LEUleuHisGlyValCysMetAlaGlyAspSerIleMetValGlnIlePheValTyrLeu 600  
 DB 1753 TTACTGACACGGGCTGCTGCTGCTGAGACACATCATGCTGACAGAAATTTTGTATCTTA 1812  
 QY 601 GLYAlaIleAspMetTyrLeuArgLysArgLysHisLeuValSerAlaSerTPHisLeu 620  
 DB 1813 GGAGCAATTGACATGATCTGCGCAAGCGTGGCCACCTGTGTACGCCAGCTGGAACCTG 1872  
 QY 621 GlnValIleThrGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis 640  
 DB 1873 CAGGTGACCAACGCTGCAATGCTCCCTTAATCTTGAGAGCAAAAGGCTTCTCTCAC 1932  
 QY 641 GlnAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGlyAsnProPro 660  
 DB 1933 GGCACAGTCTCAGCAAGGAGTGTCTCGGTCTGAGAGGGGCTGATGGGATCCACT 1992  
 QY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluIleLeuThr 680  
 DB 1993 TTCATTAAAGCTAGTATCTGCTGCTCAGTCCCACTGTGCTGAGCTGGAAATGCTCAC 2052

QY 681 ASPArgIleProTPyrValAlaProGlyCysLeuGlnGluAlaGlnThrLeuGlyLeuGlu 700  
 DB 2053 GACAGAAATACCTGGGTGGCCCCCAATGCTTCAGAGAGGCTCAGACACTGCTTGGAG 2112  
 QY 701 ALASPlySTPGLYPheGlyValAlaThrTPyrGluValPheGlnArgGlyProAlaHis 720  
 DB 2113 GCTGACAAATGGGGCTTTTGAGACCCACACAGTGGAGGTGTTCCAAAGGGAGCCCCCAG 2172  
 QY 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyIleLeuPro 740  
 DB 2173 ATCACTCGCTGAGACCCGCCAAAAGCTGAAGTTCATAGAGACAGGACAGCTGGCCC 2232  
 QY 741 ALAleuLysTPThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760  
 DB 2233 GCTTCAAATGACAGAACTGGGGACTTATCACACAGTGCATGGCTGATGATCTGGC 2292  
 QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780  
 DB 2293 CCGCGCCCTCTCTCGAGACTATCTCAGAGACCTCAAGGGCTCATTAACATCAATTAC 2352  
 QY 781 GlnLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800  
 DB 2353 GAGCTCTCTCAGACCCACACACTGCAATCCGAGTCTCTCGAGATGAGCTGTGCTAGCT 2412  
 QY 801 GLYAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLysTyr 820  
 DB 2413 GCGCGCCACGCTTATGCTGCTGCGAGACCCCGCCATATTTCAGAGAGACACCTTAAGTAC 2472  
 QY 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840  
 DB 2473 ATCTCTTGTGCTGGCAAGGGCAACTTGTGACGCGTGGAGCTGTGCGCTATGACCCCTG 2532  
 QY 841 GLYAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860  
 DB 2533 GGGGCAATACGGGACCCCTGCTGCTGAGTGAACAGCTACAGCAGCGTCCCAACACAG 2592  
 QY 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880  
 DB 2593 CAGAGGACTTCCACGCGGAGATTCAATCTTAAGGCTCTGACAGGACTTCACTCTTC 2652  
 QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValIleMetGluTyr 900  
 DB 2653 AAGTACCGGGAGTACGCTATGAGGCAAGTGCACAGACCTGCGCTTGTGATGAGTAC 2712  
 QY 901 LEUProSerGlyCysLeuArgAspLeuLeuGlnArgHisAspGlyLeuHisTPHisArg 920  
 DB 2713 TTGCCCAAGGCTGCTGCTGCTGAGACTTGTGCAAGCCCTATCCGGGCTGACACCGACCTG 2772  
 QY 921 LEUleuLeuPheAlaTPRglnIleCysLysGlyMetGluTyrLeuGlyAlaArgArgCys 940  
 DB 2773 CTACTGCTGTTCGGTTCGACATGTCGAAGGGCATGAGTACCTGGGTGGCGCGGTGC 2832  
 QY 941 VALHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHisValLysIle 960  
 DB 2833 GTACACCTGACCTGGCTGCGCGCAACATCTTGTGGAGACCGAGGCTCATAGTGAAGTTC 2892  
 QY 961 ALASPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGlu 980  
 DB 2893 GCGGACTTCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2952  
 QY 981 PROGlyGlnSerProIlePheTPyrAlaProGlnSerLeuSerAspAsnIlePheSer 1000  
 DB 2953 CCTGGCCAAAGCCCATCTTTTGGATGCCCGGAGTCCCTATGCAACATCTTCTCC 3012  
 QY 1001 ArgGlnSerAspValTPRSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAsp 1020  
 DB 3013 CGCCCAITTGACGTGGAGTTTCGGAAGTGTGTGTGACGAGCTTCCTACCTACTGCGAC 3072  
 QY 1021 LysSerCysSerProSerAlaGluPheLeuArgMetGlyProGluArgGluGlyPro 1040  
 DB 3073 AAGAGCTCAGACCCCATCGCTGAGTCTCTGGCATGATGGGGGCTGAGAGCGTGAAGACCC 3132  
 QY 1041 PROLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProProProThr 1060

|||||  
Db 3133 CCGCTGCGCCGCTGCGAGCTGTCGACAGGCGACCCCTCCACCACTCCAC 3192  
Oy 1061 CysProthrgluValGlnGluLeuMetGlnLeuGlyStrAlaProGluProHisAspArg 1080  
|||||  
Db 3193 TGGCCACCGAGGTTTCAGGAGCTCATGCTGTGGCGGCCGAGCCGACGACGG 3252  
Oy 1081 ProAlaPheAlaThrLeuSerProGlnLeuAspProLeuThrPargGlyArgProGly 1099  
|||||  
Db 3253 CCAGCTTCGCGACCTGAGCCCGACGCTGAGCCCGCTGTGGCTGGAAGACCGCGA 3309  
RESULT 2  
MUSJAK3A 3528 bp mRNA linear ROD 24-APR-1996  
LOCUS Mus musculus JAK3 gene, complete cds.  
DEFINITION LA0172.1 GI:1019909  
ACCESSION LA0172.1 GI:1019909  
VERSION JAK3 gene.  
KEYWORDS Mus musculus (strain BALB/c x 129 F2) (clone: clone-s12-18) neonate  
SOURCE thymus cDNA to mRNA.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 3528)  
AUTHORS Gurniak C.B. and Berg L.J.  
TITLE Murine JAK3 is preferentially expressed in hematopoietic tissues  
and lymphocyte precursor cells  
JOURNAL Blood 87 (8), 3151-3160 (1996)  
MEDLINE 96184772  
PUBMED 8605329  
FEATURES  
source location/Qualifiers  
1..3528  
/organism="Mus musculus"  
/strain="BALB/c x 129 F2"  
/db\_xref="taxon:10090"  
/clone="clone-s12-18"  
/cell\_type="T cell precursor"  
/tissue\_type="thymus"  
/dev\_stage="neonate"  
1..3528  
/gene="JAK3"  
1..71  
/gene="JAK3"  
72..3374  
/gene="JAK3"  
/codon\_start=1  
/protein\_id="AAC42085.1"  
/db\_xref="GI:1019910"  
/translation="MAPPSEETPLIPQSCGLSSSEAGALHVLPPRGCPQRISS  
FGDIADLCVRAKAGILPVHSLPALATEDSCMPPSHIFCIDVDVQVLYRL  
RFPDWFLETCRRKFGILKDLTSAIDLHVLHLEFQHRSDLSGLRPLVGLSKKEOG  
EFILAVLDAOMARBOAORPGLKTVSYKACLPISLRSDIIOGQENFTRRIRRTV  
LAIRRVACADARALMAKYTLDELRLHAPATPTFPVGLDPAOEEGLLRVADNGI  
SWSSGDELRFOTPCDFPEIYDVSIKQAPRGPAENHLYTYRMDGHLLEPFGLE  
ALSTVALVDGFRILCDSRHFEKVEAPRKLLEAEACHPITLDRATIKRLASGL  
PGTYILRRSPQYDSFLITACVQPIGPDPKGLIRDPSCAFSLVGLSDHRSRL  
LAACMNSGLRVGALNLTSCCAPRPEKMLIVRGCTPAPAPGSPSCALTOLS  
FHTIPTDLSLEHNLGHSFTKIPGRREVDGFTHDEYLLKVMDSRHNCESL  
EASLSMOVSYPHLYLHGYCMAGDSIMVOFEYGLAIDMLRKGLHVSWMLOYT  
KQALVALNLTEDKLPNGNSARKVLLARBEQDNPPIKISDPCVSPYVLSLEMLD  
RIPWAPBCELOEOTLCLEADKGFATWVFSGGPAHITSLPARKKLYEDOGOL  
PALMTLGLITQCMAYDPGRSPFRLILNDLGLITSDYELLSDPGLIPSPRDEL  
CGGALVACODPAPFEERHLYISLKGNGFSYELCRYPDLGNTGSLVAVKOLHS  
GPGQORPQREIOLIKALHSDFIYKRGVSYPGROSILVGMTNIPSGCLRDFQHR  
ARLTEDLLIFAMQICGMETIAGRCYHRLARNTLIVSEAHVKIADFLATLPL  
GKDIYVREPGQSPITFWAPSLSDNIFSRSDWSVSGVYLYELFTCDKCSASAP  
LRMGPEREGPPLRLLELAEGRRLLPPPTCPTEVLELMOLCWAPSHDRPARGTILS  
POLDALWGRPG"  
3' UTR 3375..3528  
/gene="JAK3"  
polyA\_signal 3517..3522  
/gene="JAK3"  
BASE COUNT 648 a 1104 c 1036 g 740 t

ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 3528  
Score: 5687.50 Matches: 1080  
Percent Similarity: 98.19% Conservative: 2  
Best Local Similarity: 98.00% Mismatches: 16  
Query Match: 97.06% Indels: 5  
DB: 10 Gaps: 3  
US-09-397-967-16 (1-1099) x MUSJAK3A (1-3528)  
Oy 1 MetalapProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
Db 72 ATGGACACTTCCAATGAGAGACACCTCTGATCCCTAGCGCTTTCACGCTTCATCC 131  
Oy 21 SerGluAlaGluAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
Db 132 TCAGAGCAGAGAGCCGTCATGCTCTCTCCCGGGGACCTGGGCTCCCGACCGA 191  
Oy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaAla 60  
Db 192 TTGTCATTCCTTTTGGGAGACTACTGCTGAGGATTTATGTGCCAGCTCCAGGCC 251  
Oy 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
Db 252 TGTGGCATCCGCTGTTTATCATTCGCTTTCGCTCGCCACTGAGACTTCTTGGC 311  
Oy 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
Db 312 TGGTTTCCCGCCAGCCACATCTTGCATAGAGAGCGGACGACATCAAGCTCTGCTAC 371  
Oy 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
Db 372 AGGCTACCTTTATTTCCCTGACTGTTGGCTGGAGATGTCACCGCTTGGGCTG 431  
Oy 121 ArgIysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
Db 432 CGCAAGATTTGACCAAGTGCATCTTGACTTACATGTTTAAACATCTCTTGTGCTAG 491  
Oy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetIysGluGly 160  
Db 492 CACGCACTGACTGAGAGAGAGGCGGCTCCCGGTGGGCTTACATGAGAGACAGGA 551  
Oy 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
Db 552 GAGTTCCGAGACCGGCGGCTGCTGACTTGGCCAGATGAGCTCGTAGACAGCGCCAGGCG 611  
Oy 181 ProGlyGluLeuLeuValSerTyrTyrAlaCysLeuProProSerLeuArgAsp 200  
Db 612 CCAGAGAGCGCTGMAAGCGGTCAAGTTCAAGCTGTCTGCGCCAGCGCTCGCGAT 671  
Oy 201 ValIleGlnGlyGlnAspPheValThrArgArgIleArgArgThrValValLeuAla 220  
Db 672 GTGATCCAGGGCCAGAACTTGTGACACAGCGGCATCCGACAGACCGTGTGGCG 731  
Oy 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaIstYrIleLe 240  
Db 732 CTGGC-CCGTGTGGCGCTGCGACGCGGACCGCTACCGCTCATGGCAAGTATTCT 790  
Oy 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValIleGlyLeuProG 260  
Db 791 GGACCTGAGGGGCTGACATCCAGCGGCGCACACGAGACCTTCGTGTGGGCTCCCGGG 850  
Oy 260 ValGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrPheSer 280  
Db 851 CCGCCAGAGAGAGCGCGGCTTCTCGTGTGGCGGAGACACGCAATCTCTGGAGCTC 910  
Oy 280 rAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerI 299  
Db 911 CGGGACACGAGAGCTTTCCAGACCTCTGTGACTTTCGGAATAACGTGGATGTCAGCAT 970  
Oy 299 eAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMe 319



Db 971 CAAGAGAGCCCGCCAGTGGGTCCGCGAGGAGCAGCGGCTGCTACGTCCACCGAGAT 1030  
 Oy 319 LAspGLyHsILeUgUaLaGluPhpProGlyLeuProGlyUaLaLeuSerPheValAl 339  
 Db 1031 GGACGGCCCACTCGAGAGCGGAGTTCGGGGCTGCGAGGCGCTGCTCTTCTGGTGC 1090  
 Oy 339 aLeuValAspGLyTrPheArgLeuILeCysAspSerArgHsTrPheCysLysGLyVa 359  
 Db 1091 CCTCGTGGATGGGTACTTCCTCCGCTGATCGCAGCTCCAGGCACTTATTTCTGCAAGAGGT 1150  
 Oy 359 lAlaProProArgLeuLeuGLuGLuAlaAspValCysHsISGLyProILeThLeuAs 379  
 Db 1151 GGCGCGCCAGCGCTGCTGAGAGAGAGCGGAGCTGTGCATGACCATCATCCTTGA 1210  
 Oy 379 pPheAlaILeHsILysLeuLysAlaAlaGLySerLeuProGlyTrPTrYLLeuArgAr 399  
 Db 1211 CTTTGCCATCCCAAGACTGAAAGCGGCTGGCTCCCTCCAGGCACTTATTTCTCGCGC 1270  
 Oy 399 gSerProGlnAspTrYAspSerPheLeuLeuThraLacysValGlnThrProLeuGLyPr 419  
 Db 1271 CAGCCCGCAGAGCATATGACAGCTTCTTCTTACGCGCTGCGTCCAGACTCCTTGGGCC 1330  
 Oy 419 oAsPTrYrLysGLyCysLeuILeArgGlnAspProSerGLyAlaPheSerLeuValGLyLe 439  
 Db 1331 CGACTACAGAGGCTCCCTCATCCGCGAGAGCCCAAGCGGCGCTTCTCCCTGTGGCT 1390  
 Oy 439 uSerGlnProHsIArgSerLeuArgGLuLeuLeuUaLaIaCysTrPAspSerGLyLeuAr 459  
 Db 1391 CAGCCAGCGCCCAAGAGCTGCGGAGTGGCTTGCAGCTGCTGGAATCTGGGCTGG 1450  
 Oy 459 gValAspGLyAlaAlaLeuTrYLLeuThrSerCysAlaProArgProLysGLyLysSe 479  
 Db 1451 AGTAGACGCTGCTGCCGTAACCTTAACATCTGCTGCGCTCCCAAGACCAAGAAAGTC 1510  
 Oy 479 rAsnLeuILeValAlaArgArgGLyCysAsnProAlaProAlaProGLyCysSerProSe 499  
 Db 1511 CAATTGTATGCTGGTGGCGAAGGGGCTGCACCCCGCGCTGCTGCTGCTGCTGCCGTC 1570  
 Oy 499 rCysAlaLeuThrGlnLeuSerPheHsIThrILeProThrAspSerLeuGLUTPrH 519  
 Db 1571 CTGCGTGGCGTGCAGACAGCTAGCTTCCACACATTCACAGAGAGAGCTGAGTGGGA 1630  
 Oy 519 sGLuAsnLeuGLyHsISGLySerPheTrhLysILePheArgGLySerArgArgGLyValVa 539  
 Db 1631 CGAGAACCTGGGCTCAGCGGCTTCTTACCAAGATCTTCCGTGCGCCAGCGGAGAGTGC 1690  
 Oy 539 lAspGLyGLuThrHsAspSerGLuValLeuLeuLysValLeuLAspSerArgHsIArgAs 559  
 Db 1691 GCATGGTAGACACATGACTCGGAAGTCTCTGGAAGTCATGAGACTCCAGACATCGGAA 1750  
 Oy 559 nCysMetGlnSerPheLeuGLuAlaLAserLeuMetSerGlnValSerTrYProHsISe 579  
 Db 1751 CTGCGATGAGCTTTTCTTGGAAAGCCGCAAGCTTGATGAGCCCAAGATTCCTACCCG 1810  
 Oy 579 uValLeuLeuHsISGLyValCysMeLalAGLyAspSerILeMetValGlnGLuPheValTr 599  
 Db 1811 GGTGTACTGCGACGCGCTGCTGATGGTGCAGACATCATGATGCGAGATTTGTGA 1870  
 Oy 599 rLeuGLyAlaILeAspMetTrYLLeuArgLysArgGLyHsISLeuValLAserAlaSerTrLy 619  
 Db 1871 TCTAGGACCAATGACATGACTCGTGGCAAGCTGGCCACCTGGTGTAGCGAGCTGGA 1930  
 Oy 619 sLeuGlnValThrLysGlnLeuAlaTrYAlaLeuAsnTrYLLeuGlnAspLysGLyLeuPr 639  
 Db 1931 ACTGAGGTGACCAAGAGCTGCGCATATGCTTAACTACTTTGGAGGACCAAGGCTTCC 1990  
 Oy 639 oHsISGLyAsnValSerAlaArgLysValLeuLeuUaLaArgGLyGLyAspGLyAsnPr 659  
 Db 1991 TCACGCGCAACGCTTCAGACAGGAAGTGTCTCGCTGCTGAGGGGAGTGAATGCC 2050  
 Oy 659 oProPheILeLysLeuSerAspProGLyValSerProThrValLeuSerLeuGLuMeLe 679

Db 2051 ACCTTTCATTAAGCTAGATGATCTGTGTACAGTCCCACTGTGCTAGCCTGGAATGCT 2110  
 Oy 679 uThrAspArgILeProTrPValAlaProGLuCysLeuGlnGLuAlaGlnThLeuCysLe 699  
 Db 2111 CAGCGACAGAAATACCTGGGTGGCCCCGAAAGTGTCCAGAGGAGTCAAGACATCTGCTT 2170  
 Oy 699 uGLuAlaAspLysTrPGLyPheGLyAlaLThrHsTrPGLuValPheGlnArgGLyProAl 719  
 Db 2171 GGAGGCTGACAAAGTGGGGTGTGGAGCCACCAACGTTGGAGGTCTTCAAGGGGGAGCCGCC 2230  
 Oy 719 aHsILeHsSerLeuGLuProAlaLysLysLeuLysPheTrYGLuAspGLyGLyLe 739  
 Db 2231 CCACATCACTCTCGTGGAGCCCGCAAAAGCTGAAGTCTATGAGGACCAAGGAGAGCT 2290  
 Oy 739 uProAlaLeuLysTrPTrHGLuLeuAlaGLyLeuILePTrGlnCysMeLalATrYAspTr 759  
 Db 2291 GCCCGCTCTCAATGAGACAGACTGGCGGACTTATACACAGTGCATGGCTATGATGCC 2350  
 Oy 759 oGLYArgArgProSerPheArgAlaILeLeuArgAspLeuAsnGLyLeuILeThrSerAs 779  
 Db 2351 TGGCGGGGCGCCCTCTCCAGAGTATCCTCAGAGACCTCAACGGGCTCATTAACATCAGA 2410  
 Oy 779 pTrYLLeuLeuLeuSerAspProThrProGLyILeProSerProArgAspGLuLeuCysVa 799  
 Db 2411 TTACGAGCTCCTCTCAGACCCCACTGGGCATCCCGAGTCTCTCGAGATGAGCTGTGC-- 2468  
 Oy 799 lAlaGLyAlaGlnLeuTrYrAlaCysGlnAspProAlaILePheGlnGLuGlnArgHsILeU 819  
 Db 2469 -GGTGGCGCCAGCTCTATGCTGCCAGAGACCCCGCATATTCGAGAGAGACACCTTAA 2527  
 Oy 819 sTrYLLeSerLeuLeuGLyLysGLyAsnPheGLySerValGlnLeuCysArgTrYAspPr 839  
 Db 2528 GTACATCTCTTGTGGGCAAGGCAACTTTGGAGCGTGGAGCTGTGCCCTATGAGACC 2587  
 Oy 839 oLeuGLyAspAsnTrHnglyProLeuValAlaValLysGlnLeuGlnHsISerValProAs 859  
 Db 2588 CCGGGGAGCAATAGGGAGCCCTGTGGCACTGAACAGCTACACAGCAGCGGGCCADA 2647  
 Oy 859 pGlnGlnArgAspPheGlnArgGLuILeGlnLeuLysAlaLeuHsISerAspPheIL 879  
 Db 2648 CCAGCAGAGGAGCTTCCAGCGGGAGATTCAGATCTTAAGCTGTGCACAGGACTTCAT 2707  
 Oy 879 eValLysTrYrArgLysValSerTrYGLyProGLyArgGlnSerLeuArgLeuValMetSL 899  
 Db 2708 CCGTCAAGTACCCGGGAGTACAGTATAGGCGCAGAGCTGCGCGCTGCTGATGGA 2767  
 Oy 899 uTrYLLeuProSerGLyCysLeuArgAspLeuLeuGlnArgHsIAsArgGLy--LeuHsIsh 918  
 Db 2768 GTACCTGCCAGCGGCTGCTGCGAGACTCTCTGCAAGCGCCATCGCGCCGCTCCACAC 2827  
 Oy 918 rAspArgLeuLeuLeuPheAlaTrPGLnILeCysLysGLyMeTGLuTrYLLeuGLyAlaAr 938  
 Db 2828 CGACCGCTACTGCTGTTCGCTTGGCAGATCTGCAAGGAGCATGAGTACTGGTGGCGG 2887  
 Oy 938 gArgCysValHsIArgAspLeuAlaAlaArgAsnILeLeuValGlnSerGLuAlaHsISVa 958  
 Db 2888 CCGTGGCTGACACCGTGAACCTGCTGGCGGCAACATCTTGGTGGAGAGGAGCGTCACT 2947  
 Oy 958 lLysILeAlaAspPheGLyLeuAlaLysLeuLeuProLeuGLyLysAspTrYTrYValVa 978  
 Db 2948 GAAAGTCCCGGACTTGGCTGCTGAACCTGCTGCGCCCTGGGAAAGGACTTACTAGTGTGT 3007  
 Oy 978 lArgGLuProGLyGlnSerProILePheTrPTrYrAlaProGlnSerLeuSerAspAsnIL 998  
 Db 3008 CCGGAGCGCTGGCCAAAGCCCATCTTTTGGATATCCCGGAGTGCCTATCTGACAAACAT 3067  
 Oy 998 ePheSerArgGlnSerAspValTrPSerPheGLyValValLeuTrYrGLuLeuPheTrHTrY 1018  
 Db 3068 CTTTCCCGCAATGTGACGTTGTGAGCTTGGAGCTTGGAGTGTGTGAGAGCTTCACTCA 3127  
 Oy 1018 rCysAspLysSerCysSerProSerAlaGLuPheLeuArgMeTMeTGLyProGLuArgGL 1038  
 Db 3128 CTGCGACAAAGAGCTGACCCCATCCGCTGAGTTCCTGTGGCATGATGAGGGGCTCGAGCGTGA 3187



QY 261 AlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSer 280  
 Db 841 GCCCAGAGAGAGCGGGTTCCTGCTGACAGGGGATTAATGGCTCCCGAGAGCTCC 900  
 QY 281 AsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIle 299  
 Db 901 AAGGACAGGAGGCTTTTCCAGACCTTCTGTGACTTCCCGAATGATGATGATCCAGCATC 960  
 QY 300 AsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMet 319  
 Db 961 AAGCAGGCTCCACCGCTGGGTCCGCGAGGAGACACCGGCTGATGACCATCACCAGATG 1020  
 QY 320 AspGlyHisIleLeuGlnAlaGluPheProGlyLeuProGluAlaLeuSerPheValAla 339  
 Db 1021 GAGGCTGACATCTGTGAAACCGGAGTCCCGGGGCTCCGAGGGCTCTCTTGTGCGG 1080  
 QY 340 LeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysGlyGluVal 359  
 Db 1081 CTCGTGATGGTACTTTTCGCTGATCTGCGACCTCCAGGCACTTCTCTGCAAGGAGGTG 1140  
 QY 360 AlaProProArgLeuLeuGluGluGlnAlaAspValCysHisGlyProIleThrLeuAsp 379  
 Db 1141 GCACCGCCACCGCTCTGAGAGAGAGCGGAGCTGTGCTCCATGACCATCAGCTTAGAC 1200  
 QY 380 PheAlaIleHisLysLeuLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArg 399  
 Db 1201 TTTGGCATTCACAGACTGAAAGCTGAGCTGCTCCCTCCCGGGCTCTTCATCTCCGACGC 1260  
 QY 400 SerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyPro 419  
 Db 1261 AGCCCGAGGACACTGACAGCTTCTTCTTACTGCTGTCTCCAGACTCTCTTGGGCCCC 1320  
 QY 420 AspyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeu 439  
 Db 1321 GACTACAGAGGCTCTCTCATCCGCCAAGACCCAGTGGGGCTTCTCCCTGGTGGCTC 1380  
 QY 440 SerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaGlyTrpAsnSerGlyLeuArg 459  
 Db 1381 AGCCAGCTGCACAGACCTTAACAGAGACTCTCTCACAGCCCTGCGGCACTTGTGACATG 1440  
 QY 460 ValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaProArgProLysGlyLysSer 479  
 Db 1441 GTAAACGCGCACTGCCCTGAACCTTCACATCTTCTGCTGCCGCCAGACCAAGGAAGTCC 1500  
 QY 480 AsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSer 499  
 Db 1501 AATTGTATCTGTCGCGTACAGGCGCCCAACCCACCGCTGCCCCGCGGCACTCCCATCC 1560  
 QY 500 CysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGlnIleTrpHis 519  
 Db 1561 TGCTGTGCGCTGACTAAGTTGAGCTTCCACACAATCCACAGACAGCTGTGAGTGGCAC 1620  
 QY 520 GlnAsnLeuGlnHisGlySerPheThrLysIlePheArgGlySerArgArgIleValVal 539  
 Db 1621 GACAATCTTGTCACGGCTCTTTTACCAAGATCTTCATAGCCATAGCGGAGGTGTGTG 1680  
 QY 540 AspGlyGlnThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsn 559  
 Db 1681 GATGGTGACACATGACACAGAAAGTCCCTGGAAGTTAAGATCCACAGACCAACAGAAC 1740  
 QY 560 CysMetGlnSerPheLeuGlnAlaIleSerLeuMetSerGlnValSerTyrProHisLeu 579  
 Db 1741 TGCATGAGACTCTTTCTGAAAGCCCAAGCTTGATGAGCAAGATATCTTACCCACACTG 1800  
 QY 580 ValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnIlePheValTyr 599  
 Db 1801 GTGTGTGCTGCATGAGTGTGATGCTGAGACAGACATCATGGTGCAGAGAAATTTGTGTAC 1860  
 QY 600 LeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLys 619  
 Db 1861 CTTGAGACAAATTGACAGCTACTGCGCAAGCGTGTCACTGCTGCGGCGACAGCTGGA 1920

QY 620 LeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspLysGlyLeuPro 639  
 Db 1921 CTCGAGTGACAAAGACAGCTGGCTGCTTAACTTACTTGGAGAGCAAAAGCCCTCCCT 1980  
 QY 640 HisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGlnGlyLysAspLysPro 659  
 Db 1981 CATGGCAACGTCACAGGAGGAAGGTGTGCTAGCTGTGAGGGGTGTATGGAGATCCA 2040  
 QY 660 ProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGlnMetLeu 679  
 Db 2041 CTTTTCATCAAGCTGAGTATCTGTGTGTCAGCCCAACTGTTTGTGAGCTTGAAATGTGCTC 2100  
 QY 680 ThrAspArgIleProTrpValAlaProGlyCysLeuGlnGlnAlaGlnIleThrLeuCysLeu 699  
 Db 2101 ACTGACAGATTCCTCGGGTGGCCCGGAGGTGTCTCCAGAAAGCTGGAACGCTCAACTTGG 2160  
 QY 700 GlnAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAla 719  
 Db 2161 GAGGCTGACAAAGTGGGCTTTTGGCCGACACAGCTGGAGGTGTTCAGCGGGGCAACCATG 2220  
 QY 720 HisIleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGlnAspGlnGlyLeu 739  
 Db 2221 CACATTAACCTCGCTGGAGCGCCCAAGACCTGAAAGTTCTATGAGAGCGGGGACAGCTG 2280  
 QY 740 ProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759  
 Db 2281 CCGGCGCTCAAAATGACAGAGCTGGAGGGCTCATCGCAGTGCGATGCGCTATGATCTCT 2340  
 QY 760 GlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779  
 Db 2341 GGCCTGGCGCCCTCTCTTCGAGACCATCTTCAGAAACCTCAACAGCGCCATTTACATCAGAT 2400  
 QY 780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVal 799  
 Db 2401 TACGAGCTCTCTGACAGCCCAACAGCTGGCATGCCGAATCCCGAGATGAGCTGTGT--- 2457  
 QY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLys 819  
 Db 2458 GGTGTGCGCCAGCTCTATGCTGTGCAGAGACCCGCCATCTTCCAGAGAGACACCTTAAG 2517  
 QY 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPro 839  
 Db 2518 TACATCTCCCTGCTGGCAAGGGGAAGTCTGGCAGCTGTGACCTGACAGACCT 2577  
 QY 840 LeuGlyAsnAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859  
 Db 2578 CTGGGACACATACGGGAGCCCTGTGTAGACGTAGACAGTACAGCAGCGGGCCAGAA 2637  
 QY 860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879  
 Db 2638 CACGACAGGAGACTCCAGGAGATTCAGATCTCTTAAGGCTCTACATCTTCACTCATC 2697  
 QY 880 ValLysTyrArgGlyValSerTyrGlyProGlyValGlnSerLeuArgLeuValMetGlu 899  
 Db 2698 GTCAAGTACCGGGGAGTACACTATGGCGCTGTGCGCCAGAGACTCGGGTTAGTGAATGGAG 2757  
 QY 900 TyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly---LeuHisThr 918  
 Db 2758 TACCTGCCACGAGCGGTGCTGGGGGACTTCTGTGACAGCCGACCGGCGGCTGCACAAAC 2817  
 QY 919 AspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlnValAlaArg 938  
 Db 2818 GACCGGCTACTGCTGTTCCTGGGAGATCTCAAGAGGCATGAGTACTGTGGGCGACAGC 2877  
 QY 939 ArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValGlnSerGlnAlaHisVal 958  
 Db 2878 CCTGCTGCATCGTACCTGTGCTGCGGCAACATCTCTGGTGAAGAGTGAAGCGCCATGTG 2937  
 QY 959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuPheLeuGlnLysAspTyrTyrValVal 978  
 Db 2938 AAGATTGCTGACTTGAGCTGGCTGCTAAGCTGTGCTTGAAGAAAGTACTACTAGCTGTG 2997  
 QY 979 ArgGluProGlyGlnSerProIlePheThrPyrIleThrProGlnSerAspAsnIle---998

```

Db 2998 CGCGTGGCGGCGCAAGTCCATCTCTGTAGCGCCCTGAGTCCCTATCTGACACATC 3057
Oy 999 pHeSerArgInserAspValTyrSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
Db 3058 TTTCTCCGCCCATCTGACGCTGTGAGCTTGGCGGTGTCTGTATGAGCTCTTCCACTAC 3117
Oy 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu 1038
Db 3118 AGCAGCAAGACCTGCACGCCATCCATGAGTCTCTGCTATGATGAGCGCTGAGCGAG 3177
Oy 1039 GlyProProLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProProPro 1058
Db 3178 GGATCCCGGCTCTGCACCTTGTGAGCTGTGCGACAGGCGCCGCTCTCCACACGCC 3237
Oy 1059 ProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTTPAlaProGluProHis 1078
Db 3238 TCCACCTCCCTACTGAGCTTCAAGAGCTCATGAGCTGTGCTGTGCGCCACCCACCA 3297
Oy 1079 AspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTyrArgGlyArgPro 1098
Db 3298 GACGCGCCAGCTTTGACACACCTTGAGCCGCCAGCTGAGTGGCGCGGCGGAGGCC 3357
Oy 1099 Gly 1099
Db 3358 GGA 3360

RESULT 4
MUSJAK3H 4016 bp mRNA linear ROD 13-SEP-1994
LOCUS Mus musculus (clone 32D5-1) protein tyrosine kinase (JAK3) mRNA.
DEFINITION L33768
ACCESSION L33768.1 GI:508533
VERSION protein tyrosine kinase.
KEYWORDS Mus musculus cDNA to mRNA.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
AUTHORS Rane,S.G. and Reddy,E.P.
TITLE JAK3: a novel JAK kinase associated with terminal differentiation
of hematopoietic cells
JOURNAL Oncogene 9 (8), 2415-2423 (1994)
MEDLINE 94309920
PUBMED 7518579
FEATURES
source
location/Qualifiers
1..4016
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="32D5-1"
/cell_line="32Dc13"
/tissue_lib="lambda gtl1"
7..3906
/note="protein tyrosine kinase"
/codon_start=1
/product="JAK3"
/db_xref="GI:537607"
/translation="MAPSSEPPPLIPORSCLSSSEAGALHVLPPRGPGPPORTSFS
FGDYLADLVRAKACALIPYHSLPALTEPFSCMPSPHFLIEDVQVLYRL
RFYFPMFGLTCHREGLRDLTSALIDLVEHLRAHRSIDVSGRLVSMKSG
EFLSLAVDLAQAAREQAPRGELTIVSKACLPRLVDTGQNFYRRRTIRPVY
LALWVSPARTATAGOVYLDLERLPAATETFRVGLPGAEPEGGLRAGDNCSIS
WSSGDEVILGILRSVGRGAMRLVKPVRTIRKVASPGRADCTTGQGGVNLKVPG
MEDPGITWCTRVRLDRGRTEGSDMWISGHDTPRVSPLTSSPPHKKWEG
GRRGCGAKSRVITWLSLLEFPNFCDFCPPTVYVSTICPGSGRHRITVTRM
CLSSAVIKGRVPLRYCLSELPOLFTFCDFPRTVYVSTICPGSGRHRITVTRM
DGHILEFPELPALSFVALVDYFRLTCDSTRTYRKEVAAARLLEERALECHGPT
LGASGQOMGMGAARTVLGWTWLSMRPLDFAIHKAKAASLREGTILRSQDSDS
FLTACVOTPLGPDKLIGLIDPSGAFSLVSAASPEACGSCILDERLKDGAAL
NLTSCAPRPEKNLIVRGCCTPAPAGCSPSCALITDLEPHITPTDVLGHENLGH
GSFTKIRRRRRREVDETHSEVLLVKVMSRHRNCESFLEAASLMSQVSYHLVLL
HGVCAGDSINWQEVYLGALIDMLRKRGHLVSAWMLQYTKOLAVLALNLEDKGLPH

```

```

GNSARKVLLAREGGDGNPEFTKIDPGVSPVLSLEMLTRIPWVAPECLQEAOTLC
LEADKMGEGATTEVEVSGGPAHITSLPEAKKLEFEDOGALMKMTLGLITQMA
YDPGRSPFRPAIILRLDNLITSDYELLQTPHPLASRLVCLGPRHTEER
HLKXISLLGKNGSVELCRYPDLMDTGLVAKQLOHSGPROOGRFOREIOLKALH
SDITVAKRGVSYSPGROSLRLVMEYLPSSGLRPFQDRHRAALTDRLLEFMOITKM
EYIGARCVHRLAARNILIVESEAHVKIADFGIARLLPLGKDYVREGGSPITWYA
PEISDNIIFRSQSDVMSFGVLYLELFTYCDKSCSPSAEFLRMGPRRESPLCRLEL
LABGRRLPPEPTCTEYQELMOCLVAPSRTTQGPAP"
2199..3024
/note="JAK homology domain II (JH2)"
misc_feature
misc_feature
BASE COUNT 742 a 1203 c 1235 g 836 t
ORIGIN
Alignment Scores:
pred. No.: 1,17e-311 Length: 4016
Score: 5223.00 Matches: 1060
Percent Similarity: 80.59% Conservative: 3
Best Local Similarity: 80.36% Mismatches: 32
Query Match: 89.13% Indels: 235
Gaps: 6
DB: 10
US-09-397-967-16 (1-1099) x MUSJAK3H (1-4016)
Oy 1 MetalProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 7 ATGACACCTCCAGTAGAGAGACACCTCGATCCCTGACGCTGTGACGCTCTCATCC 66
Oy 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40
Db 67 TCAGAGCAGAGAGACCTCGATGTGCTCTTCCCTCCCGGAGACCTGGCGCTCCACAGCA 126
Oy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValAlaArgAlaAlaVal 60
Db 127 TTGCATTTCTCTTTGGGAGCTACTTGGCGAGATTATGTGTGGAGCTGCCAAGGCC 186
Oy 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 187 TGTGCGATCGCTGCTGTTCATTCGCTTTGCGCTGGCCACAGGACCTTCTCTTGC 246
Oy 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 247 TGGTTTCCCGCCACACCATCTTGCATGAGAGACCTGACACTCAAGTCTTGCTTAC 306
Oy 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 307 AGGCTACGCTTTATTTCTCCGACTGTGTGGCTGGAGACATGTACCGCTTGGGCTG 366
Oy 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 367 CCCAAGATTGACACAGTGCATCTTGACTTACATGTTTAGACATCTCTTGTCTCAG 426
Oy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
Db 427 CACCCAGTAGACTGTGAGTGGGCGCTCCCGGTGGGCTTACCATGAAGACAGGCA 486
Oy 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArg 180
Db 487 GAGTTCTTAGCCTGGCGCTGTGACTTGGCGCCAGATGCTCGTAGACAGCCACAGC 546
Oy 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
Db 547 CCAGGAGACCTGCTGAAGCGCTCAGTTACAAAGCTGTGCGCCCGAGCTGCGCAT 606
Oy 201 ValIleGlnGlyGlnAspPheValThrArgArgGlyIleArgArgThrValValLeuAla 220
Db 607 GTATCCAGGCGCAGACTTCGTGACACCGACCGCATCCGAGGACCGTGTGTGGCG 666
Oy 221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240
Db 667 CTGC--GCGTGTGCTGCGCTGCAGGCGCAGCGCTACGCTCATGCGCCAGATATAT-CTG 723

```



Db 2875 GCTCTCAAAATGAGACAACTGGCGGACTTATCACACAGTGCATGGCGTATGATCTGGC 2934  
 QY 761 ArgArgProSerPheArgAlaIleLeuAArgPheuAsnGlyLeuIleThrSerAspTyr 780  
 Db 2935 CGGGGCCCCCTCTCCAGCTATCTCTCAGACACTCAAGCCCTCATCTACATCGATTAC 2994  
 QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800  
 Db 2995 GAGGCTCCCA-GACCCACACCTGGCATCCCGAGTCCCTCAGATGAGCTGTGGT---- 3049  
 QY 801 GlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluAAsnHisLeuIleTyr 820  
 Db 3050 GGGCCCGACTATGCTGGCAGACCCCGCATATC-GAGGAGAGACACCTTAAGTAC 3108  
 QY 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840  
 Db 3109 ATCTTTTGCTGGCAGAGGCACTTTGGCAGCGTGGAGCTGTGGCTATGACCCCTGG 3168  
 QY 841 GlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860  
 Db 3169 ---GACAAATACGGGACCCCTGTGGCTAGTGAACAGCTACAGCAGCGGGCCAGAC 3225  
 QY 861 GlnArgAspPheGlnArgGlnIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880  
 Db 3226 CAGAGGAGCTTCCAGCGGAGATTCAGATCCTTAAGCTGTCCACAGGACTTCATCGTC 3285  
 QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyr 900  
 Db 3286 AAGTACCGGGAGTACAGCTATGGGCGGAGTCCCGAGGCTGGGTGGTGGAGTAC 3345  
 QY 901 LeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHisThrAsp 919  
 Db 3346 CTGGCCAGGGGCTGCTGGAGACTTCTCGACGCCCATCGCGGGCCCTGCACACCGAC 3405  
 QY 920 ArgLeuLeuLeuPheAlaIlePheGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg 939  
 Db 3406 CGCTACTCTGTGCTGGCAGATCTCAAGGCGATGAGTACCTGGTGGTGGCGGCC 3465  
 QY 940 CysValHisArgAspLeuAlaIleArgAsnIleLeuValGlnSerGlnAlaHisValLys 959  
 Db 3466 TGGTACACCGTACCTGCTGGCGCAACATCTTGGTGGAGGCGAGGCTATGTGAAG 3525  
 QY 960 IleAlaAspPheGlyLeuAlaIleLeuLeuProLeuGlyLysAspTyrTyrValAlaArg 979  
 Db 3526 ATGGCGAGCTTGGCTCGGTAGCTGCTGGCCCTGGAAAGACTACTACGCTGCC 3585  
 QY 980 GluProGlyGlnSerProIlePheTyrTyrAlaProGlySerLeuSerAspAsnIlePhe 999  
 Db 3586 GAGCCTGGCCAAAGCCCATCTTTGGTATGCCCCGAGTCCCTATCTGACACATCTTC 3645  
 QY 1000 SerArgGlnSerAspValTrpSerPheGlyValIleLeuTyrGlnLeuPheThrTyrCys 1019  
 Db 3646 TCCCGGCAATCTACGTGTGAGCTTCGAGTGTGTAGCAGCTCTTACCTACCTGCG 3705  
 QY 1020 AspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGly 1039  
 Db 3706 GACAAGAGCTGACAGCCATCCGCTGAGTCTCGCATGTGGGCGCTGAGCGCTGAAGA 3765  
 QY 1040 ProProLeuLeuArgLeuGlnLeuLeuAlaGlnGluIleArgTyrLeuProProPro 1059  
 Db 3766 CCCCCGCTGGCGCTCTCTGGAGCTGTGGCAGAGGGCGGAGCGCTCCACACCACTCC 3825  
 QY 1060 ThrCysProThrGlnValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHisAsp 1079  
 Db 3826 ACCTGCCCCACGAGGTTCAGAGCTCATGACGCTGTGGCGGCCAG-CCGACAGAC 3884  
 QY 1080 ArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgPro 1098  
 Db 3885 CGGCACACCTTCGACACCTGAGGCCCGCCAGCTGAGCGCTGTGGCTGGAAGACC 3941  
 RESULT 5  
 AX203081 . 4064 bp DNA linear PAT 30-AUG-2001  
 LOCUS

DEFINITION Sequence 2 from Patent WO0152892.  
 ACCESSION AX203081  
 VERSION AX203081.1 GI:15392438  
 KEYWORDS  
 SOURCE  
 ORGANISM human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 4064)  
 AUTHORS Vastou, G.  
 TITLE Jak/stat pathway inhibitors and the uses thereof  
 JOURNAL Patent: WO 0152892-A 2 26-JUL-2001;  
 GENZYME CORPORATION (US)  
 FEATURES  
 source 1..4064  
 location/Qualifiers  
 BASE COUNT 746 a 1292 c 1168 g 858 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1 57e-278 Length: 4064  
 Score: 4684.00 Matches: 905  
 Percent Similarity: 87.61% Conservative: 64  
 Best Local Similarity: 81.83% Mismatches: 129  
 Query Match: 79.93% Indels: 9  
 Gaps: 5  
 DB: 6  
 US-09-397-967-16 (1-1099) x AX203081 (1-4064)  
 QY 1 MetaLapProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSer 20  
 Db 96 ATGGACCTCCAAATGAGACAGCCCTGATCCCTCAGGCTTCATCAGCTCTGTGC 155  
 QY 21 SerGlnAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40  
 Db 156 ACGAGGCTGTGTGCTGCATGTGCTGCTGCCCGCGGGGCCCGGCCACAGCG 215  
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValAlaAlaValAla 60  
 Db 216 CTATCTTCTCTCTTGGGAGGACTGTGCTGAGGAGCTGTGCTGAGGCTGCCAAGGCC 275  
 QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
 Db 276 AGCGCATCTGCTGCTGTACACTCCCTCTTGTGCTGCGCAGGAGGACTGCTCTGC 335  
 QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
 Db 336 TGGTCCCCCGAGCACATCTTCTCCGTGAGGATGCCACACCAAGTCTCTGTAC 395  
 QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluIleThrCysHisArgPheGlyLeu 120  
 Db 396 AGGATTCCTTTTACTCTCCCAATGTGTTGGCTGGAGAAAGTGCACCGCTGGGGCTA 455  
 QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaGln 140  
 Db 456 CCCAAGATTGGCCAGATGCTATCTTACCTTGACCTGGCAGCTGGAGCACCTCTTGGCCAG 515  
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGlnGly 160  
 Db 516 CACCGCAGTACTGTGTGAGTGGGCGCCCTCCCGTGGGCTCAGTCTCAAGGAGAGGCT 575  
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArg 180  
 Db 576 GAGTGTCTAGCTGTGGCTGTGTGACCTGGCGCGGATGGCGGAGAGAGCCAGCGG 635  
 QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200  
 Db 636 CGGGAGAGCTGTGTAAGACTGCTACAGCTCAAGGCTGCTGCCCAAGCTTGGCCAG 695  
 QY 201 ValIleGlnGlyLysAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220  
 Db 696 CTGATCCAGGGCGCTGAGCTTGTGAGCGGAGGAGGCTATTTCGAGAGAGGCTGGCAGAGCC 755

```

OY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrlAlaLeuMetAlaIleuTyrlIle 240
Db 756 CTGCC-GCGCGTGGCGGCTGCCAGGACCGGACCTGCTCATGGCCAGATGATCAT 814
OY 240 uAspleuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
Db 815 GACACTGGAGGGCTGGATGCCAGCGGGGGCCGAGACCTTCCACCTGGGGCTCCCTGG 874
OY 260 yAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277
Db 875 GGGCTTGGTGGCCAGCAGCGGCTGGGCTGCTCCGGTGGCTGGTGGAGCGCGCATCCG 934
OY 277 oTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGlyuIleValAs 296
Db 935 CTGGAGCCAGGAGAACAGAGAGGCTCCAGCCCTTCTGGCATCTTCCAGAAATGCTAGA 994
OY 296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa 316
Db 995 CATAGCATCAACAGAGCGCCCGCGGTGGCCCGGAGAGACCGGCTGTGCTACTGT 1054
OY 316 lTrhArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGlyuAlaLeuSe 336
Db 1055 TACAGAGACAGACACACAGATTTTAAAGGCCAGATTCCAGGGCTGCCGAGGCTCTGC 1114
OY 336 rPheValAlaLeuValAspGlyTyrlPheArgLeuIleCysAspSerArgHisTyrlPheCy 356
Db 1115 GTTGTGGGGCTGTGGAGCGGCTTCCGGCTGACACAGGACTCCAGCACTTCTCTGTG 1174
OY 356 sLySGluValAlaProProArgLeuGluGluGluAlaAspValCysHisGlyProGly 376
Db 1175 CAAGAGGTGGACCGCGCGAGGCTGTGGAGAAAGTGGCCAGAGTGGCCAGCGGCCCAT 1234
OY 376 ePthrLeuAspPheAlaIleHisLyLeuLyAlaAlaGlySerLeuProGlyThrTyrl 396
Db 1235 CACTGTGACTTGGCATCAACAAGCTCAACAGCTGGGGGCTCAAGCTGTGCTCTCATGT 1294
OY 396 eLeuArgArgSerProGlnAspTyrlAspSerPheLeuLeuThrAlaCysValGlnThrPr 416
Db 1295 TCTCCGGCGGACCGCCGAGCACTTACAGCTTCTCTCTCACTGTCTGTCCAGAACCC 1354
OY 416 oLeuGlyProAspTyrlTyrlGlyCysLeuIleArgGluLeuLeuAlaAlaCysTrpAsnSe 436
Db 1355 CTTGGTCTCGATTATTAAGGGCTGCTCATCCGGCCAGCGCCCAAGGAACCTTCTCTCT 1414
OY 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSe 456
Db 1415 GATTGGCTCAGCGGACCGCCAGCAAGTCTTCAGAGCTCTGGCAACCTGCTGGGATGG 1474
OY 456 rGlyLeuArgValAspGlyAlaAlaLeuTyrlLeuThrSerCysAspAlaProArgProLy 476
Db 1475 GGGGCTGCACGTAGATGGGGTGGCACTGACCTCACTTCTCTGTATCCCGAGACCCA 1534
OY 476 sGluLySerAsnLeuIleValAlaArgArgGlyCysAsnProAlaProAlaProGlyCy 496
Db 1535 AGAAAGTCCAACTGATGCTGCTGCTCAGAGAGTCAACAGATTCCTCTGTGACAG 1594
OY 496 sSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProHisAspSe 515
Db 1595 TCAGCCCAATCCAAATACACACTGATGATGATGATGATGATGATGATGATGATGATG 1654
OY 515 rLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLyHisIlePheArgGlySerAr 535
Db 1655 CTGGAGTGGCATGAGAACTGGGCGCATGGGCTCTTCAACAAATTTAACGGGGGCTGTGC 1714
OY 535 gArgGluValAlaAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555
Db 1715 CCATGAGGTGGTGGATGGGAGGCCGGAAGACAGAGAGTCTCTGTAAGGTCTATGATGC 1774
OY 555 rArgHisArgAsnProCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575
Db 1775 CAAGCACAAAGACTGATGAGTCAATTCCTGGAACAGGAGCTTGATGAGCAAGTGTCT 1834

```

---

```

OY 575 rTyrlProHisLeuValLeuLeuHisGlyValAlaCysMetAlaGlyAspSerIleMetValG 595
Db 1835 GTACCGGATCTCTGTGCTGTCCACGGCGGTGTGATGCTGGAGAACGCCACCTGGTGTGA 1894
OY 595 nGluPheValTyrlLeuGlyAlaAlaIleAspMetTyrlLeuArgLyAspGlyHisLeuValSe 615
Db 1895 GGAATTTTACACCTGGGGGCGCATGACATGTATCTCCGAAACAGTGGCCACTGGTGTGC 1954
OY 615 rAlaSerTrpLyLeuGlnValThrLyGlnLeuAlaTyrlAlaLeuAsnTyrlLeuGlnAs 635
Db 1955 AGCCAGCTGGAAAGCTGCAGGTGTCAAAACAGTGGCTACGCCCTCACTATCTGAGAGA 2014
OY 635 pLySGlyLeuProHisGlyAsnValSerAlaArgLyValLeuLeuAlaArgGluGlyG 655
Db 2015 CAAGAGCTGCCCATATGCGAATCTCTCCGGAAAGGTGTCTGTGGCTCGGGAGGGGGC 2074
OY 655 yAspGlyAspProPheIleLyLeuSerAspProGlyValAspProThrValLeuSe 675
Db 2075 TGAATGGAGCCCGCTTCATCAAGCTGATGACCTGGGGTCAACCGCGCTGTGTAA 2134
OY 675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGlyCysLeuGlnGluAlaG 695
Db 2135 CTTGAGATGCTCACCGACAGAGATCCCTGGGTGGCCCGCAGATGTCTCCGGAGGGCGCA 2194
OY 695 nThrLeuCysLeuGluAlaAspLyTrpGlyPheGlyValAlaThrThrProGluValPheG 715
Db 2195 GACACTTACCTTGGAAAGCTGACAAAGTGGGCTTGGCGGCCACGGCTGTGGAAAGTGTAG 2254
OY 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLyLyLeuLyAspPheTyrlGluAs 735
Db 2255 TGGCTCAACCATGCCCATCAAGTCCATGATTCGTCAAAACCTCCAAATTTATAGGA 2314
OY 735 pGlnGlyGlnLeuProAlaLeuLyTrpThrGluLeuAlaGlyLeuIleThrIncysMe 755
Db 2315 CCGGACAGAGCTGCCGGCCCGCCAGTAGGACAGAGCTGGCTCTGATTTCAACAGTGCAT 2374
OY 755 fAlaTyrlAspProGlyArgArgProSerPheAlaAlaIleLeuArgAspLeuAsnGly 775
Db 2375 GGCTATAGAGCGGCTCCAGAGGCGCTCTTCCAGCGCTCATTCGTGACCTCAATAGCTT 2434
OY 775 uIleThrSerAspTyrlGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795
Db 2435 CATCTTTCAGACTATGAGCTCTTCACAGCCCAACCTGTGCTGGCCACCTGTGA 2494
OY 795 pGluLeuCyValAlaGlyAlaGlnLeuTyrlAlaCysGlnAspProAlaIlePheGluG 815
Db 2495 TGGGCTGTGG---AATGGTCCAGGCTGTATGCTTGGCAAGACCCGACGATCTTGGAGA 2551
OY 815 uArgHisLeuLyTyrlSerLeuLeuGlyLyAspAsnPheGlySerValGlyLeuCy 835
Db 2552 GAGACACCTCAAGTACATCTCACAGCTGGGCAAGGCAACTTTGGCACTGGAGGCTGTG 2611
OY 835 sArgTyrlAspProLeuGlyAspAsnThrGlyProLeuValAlaValLyGlnLeuGlnI 855
Db 2612 CCGGTATGACCGGCTAGGCCACAAATACAGGCTCTGTGGCCCTGAAACAGCTGCAGCA 2671
OY 855 sSerValProAspGlnGlnArgAspPheGlnArgGlyuIleGlnIleLeuLyAlaLeuH 875
Db 2672 CAGCGGCCAGACCGCAGAGGAGACTTTCAGCGGAGATTCAGATCTTCAAAACCTGTGA 2731
OY 875 sSerAspPheIleValLyTyrlArgGlyValSerTyrlProGlyArgGlnSerLeuArg 895
Db 2732 CAGTATTTTCATTTGCAGATGATGCTGTGTACGTATAGGCCCGGGCCGAGAGCTGGG 2791
OY 895 gLeuValMetGlyuTyrlLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgG 915
Db 2792 GCTGTATGAGAGTACCTGCCAGAGGCTGTGCGCGACTTCTGTCAACGCGCACCGGCC 2851
OY 915 y---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLyGlyMetGly 934
Db 2852 GCGGCTTGATGCCAGCGGCTCTTCTTATTCCTGCGAGATCTGCAGAGGCGATGAGATTA 2911
OY 934 rLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlySe 954

```



```

/protein_id="AAAI3626.1"
/db_xref="GI:508731"
/translation="MAPSEEMPLLPOKSCSLSTSEAGALVLLPAPAPPOPLSF
/Trans lation="MAPSEEMPLLPOKSCSLSTSEAGALVLLPAPAPPOPLSF
FGHLEADLVCQAKARASGILPYHSLFPLATEDLSLCPSPSHIFSVEDASTOVLNFI
RFEYPMWFGLEKCRHFRGLKRLDASLIDLRYLENLPHOHRSDLVSGRLPYGLSLKEGG
ECSLVAIIDLARMAEOAPRGELLKTSYKACLPSPSLRDLGGISLGRGLAEVADG
RALPRAVACQADNRHSLMAYIMLEKRLDPAAGAEFFHRYLYVATYRDNQITLAEFGD
GGIAMQOEVEVLDPCEDEPEIVDISIKAPRVLLLEVAEQCHPTITDPAINKLKG
LPLASLVALVDGEFLTDSQHFCEKVCVONPLGPDKCLLRSTGFFLLVGLSPHSIS
GSPRSYVLRSPDDSEFLTVCVONPLGPDKCLLRSTGFFLLVGLSPHSIS
SALTQWIDGQIVNDGVATLITSCIPREKSNLIVQRGHSPTSLVPOSQYOL
SOMTEHTRPADSLMEHNIHGHSFTKTYGCHHEVDNEDYALDKRLKRNKVPASW
LSQWLSLMSQYSTRKLYILHNGVSNARKYALREGDQSPRTIKSDGVSPLVSLSE
LQVVKOLAYALNLEDGRLHNGVSNARKYALREGDQSPRTIKSDGVSPLVSLSE
MLTRPIWVAPECIREAQTSLSEADKMPGVSRAVTRIDNLSISDYLSDPRLAVK
ROOLPAKTEFLALLIQQMAYPEVQPSRAVTRIDNLSISDYLSDPRLAVK
RDGANGKOLYACODPTIFEEHRLKYISQIGKNFSGVSLCRLYMETLIPSCILDP
LOHSGPDQDFOREIOILKALHSDPIVYKRGVSPGAPRELALYALPAKLFUE
QHRNRADASRLILYSQIGMEYLGRRCYHRIADANILVESHAKIADRELA
LILDKDQIVYREKQSPITFWVAPELSQNIIFROSDVSGFVGLIETLTYDKCS
SAELFRMGEKRDYVPLCSRLLELLEEGORLPAPRPAEYHELMKICMAPSPQDRPS
SALGPQIDMLMSGSGEGTALFTNHPKRNHLSLFS"
BASE COUNT      746 a      1292 c      1168 g      858 t
ORIGIN
Alignment Scores:      1,57e-278      Length:      4064
Pred. No.:      4684,00      Matches:      905
Score:      87,618      Conservative:      64
Percent Similarity:      81,838      Mismatches:      129
Best Local Similarity:      79,938      Indels:      9
Query Match:      9      Gaps:      5
DB:
US-09-397-967-16 (1-1099) x H5U09607 (1-4064)
Qy      1      Metalaproproserygluuthrproleuileproglinarsercysleuenser      20
      |||||||
      96      ATGGACCTCCAAAGTGAAGAACCCCTGATCCCTCAAGCTTCAAGCTGCTTTGCTC      155
      |||||||
      21      Seruinalagialaleuuhisvalleuunproproagsglyproglyproproglinar      40
      :::::
      156      ACGGAGGCTGGGCTGCATGATCTCTGCTGCCGCTGGGCCCCGGGGCCCCCAGGCG      215
      |||||||
      41      leuserpheserphesglyasprtyleuadlagluaspreucysvalargalaialysala      60
      :::::
      216      GATCTTTCTCTTGGGACCACTGGCTAGAGCACTGTGCGGTGACAGGCTGCAGAGGCC      275
      |||||||
      61      Cysglyileuuprovaltyrhisserleuuhelaleuulathrighuasppheser      80
      :::::
      276      AGCGGATCTCTGATGACACACCCCTTTGCTGTGCGACGAGGAGCACTGTCTGCG      335
      |||||||
      81      Thrpheroptoserhisilephesysileguasprvalasprthrinalvalleyalt      100
      |||||||
      336      TGGTTCCTCCCGAGCCACATCTTCTCCGGGAGGATGCCACACCCAAAGTCCTCTGAC      395
      |||||||
      101      Argleuargphetyrphedroasprtrphegiyleugithrincysnhisargphesyleu      120
      ||:::
      396      AGGATGCGTTTACTTCCCAATGGTTGGGCTGGAGAACTGCCACCGCTTGGGGCTA      455
      ||:::
      121      Arglyaspreuhserserialaleuasprleuuhisvalleugluhisleuuhelaglin      140
      |||||||
      456      CGCAAGGATTTGGCCAGCTGATCTTGAACGTGCCAGTCCTGGACACCTCTTGGCCAG      515
      |||||||
      141      Hisargseraspreuvalserglyargleuprovalglyleusermetelyuglinugly      160
      |||||||
      516      CACGCGATGACCTGATGATGGGCGGCTCCCGCTGGGCTCAAGTCTCAAGGAGGAGGT      575
      |||||||
      161      Glupheleuserleuavalaleuaspheleuaglinmetalaargluglinaglinarg      180
      |||
      576      GAGTGTCTACACCTGGCGGCTTGGACCTGGCCGATGGCGGAGAGAGGCCAGCGGG      635
      |||
      181      Proglygluleuuleuyslythrvalsettyrlylsalcysleuoproptoserleuargasp      200
      |||||||

```



Db 636 CCGGAGAGCTGCTGAGACTGTCTACAGACAGCCCTGCTACCCCAAGCCGCGGAC 695  
 Oy 201 ValIeGInGlyGlnAsnPhenValThrArgArgArgIleAArgGthValValLeuAla 220  
 Db 696 CTGATCCAGAGCCCTGAGCTTGTGACGCGAGGCGCTATTCGAGAGCGGTGCGCAGAGCC 755  
 Oy 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaIleTyrIleLe 240  
 Db 756 CTGCC-GCGCGTGGCCGCTCCAGCGACCGGACCTGCTATGGCCAAAGTACATCAT 814  
 Oy 240 uAspLeuGlnArgLeuHisProAlaAlaThrThrGlnThrPheArgValGlyLeuProG 260  
 Db 815 GGACCTGAGAGCGGCTGGATCCAGCGCGGCGCGAGACCTTCACACGTCGCGCTCCCTCG 874  
 Oy 260 yAla-----GInGInGInProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277  
 Db 875 GGGCCCTGTGTGGCCAGCAGCGGCTGGGGTGTCTCGGCTGGTGTGTACCGCGGCATCGC 934  
 Oy 277 oTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAs 296  
 Db 935 CTGGACCCAGGAGACAGAGAGGTCTCCAGCCCTTCTGCGACTTTCAGAAATCTGAGA 994  
 Oy 296 pValSerIleAsnGlnAlaProArgValAlaGlyProAlaGlyGlnHisArgLeuValThrVa 316  
 Db 995 CATTAAGCATCAAGCAGGCGCGCGCGCTGGCGCGCGGAGACACCGCTGGTGTCACTGT 1054  
 Oy 316 lThArgMetAspGlyHisIleLeuGlnAlaGluPheProGlyLeuProGluAlaLeuSe 336  
 Db 1055 TACCAGACACACACACCAATTTTAGAGCGCGAGTTCCAGGGCTGCGCGAGCTCTGTGC 1114  
 Oy 336 rPheValAlaLeuValAspGlyThrPheArgLeuIleCysAspSerArgHisTyrPheCy 356  
 Db 1115 GTTCGTGGCGCTGTGAGCGGCTACTTCCGGCTGACACAGGACTCCAGACTTCTTCTG 1174  
 Oy 356 sLySGluValAlaProProArgLeuLeuGlnGlnAlaAspValCysHisGlyProI 376  
 Db 1175 CAAGGAGGTGGACCCCGCGGCTGGAGAAAGTGGCGGACCAATGGCAGCGGCCCAT 1234  
 Oy 376 eThrLeuAspPheAlaIleHisIleLeuLeuAspAlaGlySerLeuProGlyThrTyrI 396  
 Db 1235 CACTCTGAGACTTGGCATCAACAAGCTCAAGAGCTGGGGCTCACGTCGGTCTCATGT 1294  
 Oy 396 eLeuArgArgSerProGlnAspTyrArgSerPheLeuLeuThrAlaCysValAlaGlnThrPr 416  
 Db 1295 TCTCCCGCCGAGCCCGCAGAGACTTTCAGAGCTTCCTCCACGTGTGTGTCCAGAACCC 1354  
 Oy 416 oLeuGlyProAspTyrTyrGlyGlyCysLeuLeuArgGlnAspProSerGlyAlaPheSerIe 436  
 Db 1355 CCTTGTCTCTGATTATAGGGGTGCTCTCATCGGGCGACCCACAGAGAACCTTCTCT 1414  
 Oy 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSe 456  
 Db 1415 GGTGGGCTCAGAGCCGACAGAGAGTTCGAGAGCTCTCGGACCACTGTGGGATGG 1474  
 Oy 456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaProArgProLy 476  
 Db 1475 GGGGCTGCAGCTAGATGGGTGTGAGTGCACCTCACTTCCTGTATCCCAAGCCCAA 1534  
 Oy 476 sGluTyrSerAsnLeuIleValAlaArgGlyCysAsnProAlaProAlaProGlyCy 496  
 Db 1535 AGAAAGTCCAAACCTGATCTGTGTCTCAGAGAGGTACAGCCACCCACATCATCTTGGT 1594  
 Oy 496 sSerPro---SerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515  
 Db 1595 TCAGGCCCAATCCCAATACAGAGCTGATGACATTTTCAACAAGATCCCTGTGACAG 1654  
 Oy 515 rLeuGlnThrPheHisGlnAsnLeuGlnHisGlySerPheThrLysIlePheArgGlySerAr 535  
 Db 1655 CCTGAGAGGTGATGAGAACCTGGGCGAATGGCTCTTCCACCAAGATTTAACCGGGGCTGCG 1714  
 Oy 535 gArgGluValAlaAspGlyGlnThrHisAspSerGluValLeuLeuLeuValMetAspSe 555  
 Db 1715 CCATGAGGTGGTGGATGGGAGGCGCCGAACAGACAGAGTGTCTGTGAAGGTATGATGTC 1774

Oy 555 rArgHisArgAsnCysMetGluSerPheLeuGlnAlaAlaSerLeuMetSerGlnValSe 575  
 Db 1775 CAAGCAACAAGACTGCATGAGTATCTCTCGAAGAGAGGAGCTGTATGAGCAAGTGTG 1834  
 Oy 575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValG 595  
 Db 1835 GTACCGGCAATCTGTCTGTCTCCAGCGCGTGTGCATGCTGTGGAGACAGCAGCATGGTGA 1894  
 Oy 595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615  
 Db 1895 GGAATTTGTACACTGCTGGGGCCCATAGACATGATGTGCGAAACGTTGGCCACTGTGTC 1954  
 Oy 615 rAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAs 635  
 Db 1955 AGCCACTGTGAAGCTGACAGGTGTGTAACAGCTGGCGCTACGCCCTCAACTATCTGAGAG 2014  
 Oy 635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGlyGly 655  
 Db 2015 CAAGGCGCTGCCCATGCTGCAATGTCTGTGCGGAAAGTGTCTGTGGCGGAGGGGCG 2074  
 Oy 655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675  
 Db 2075 TGATGGAGACCGCGCTTCATCAAGCTGATGACCTGGGGTCCAGCCCGCTGTGTAG 2134  
 Oy 675 rLeuGlnMetLeuThrAspArgIleProTyrValAlaProGlyCysLeuGlnGlnAlaG 695  
 Db 2135 CTTGGAGATGCTCACCGAGAGATCCCTGGGTGGCGCCCGCAGTGTCTCCGGAGGCGCA 2194  
 Oy 695 nThrLeuCysLeuGlnAlaAspLysTrpGlyPheGlyAlaThrThrTrpGlyValAlaPheG 715  
 Db 2195 GACACTTAGCTTAGGAAGTGCACAGATGGGGCTTCGGGCCACCGTCTGGGAAGTGTAG 2254  
 Oy 715 nArgGlyProAlaHisIleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGluAs 735  
 Db 2255 TGGCTCACCATGCCCATCACTGCTGCTGATCTGTAAAGAACTCCAAATTTATAGAGA 2314  
 Oy 735 pGlnGlyLeuLeuProAlaLeuLysTrpThrGlnLeuAlaGlyLeuThrGlnCysMe 755  
 Db 2315 CCGGAGCAGCTGCCCGGCGCCCAAGTGCAGACAGACTGGCGCTGTGATCAACAGTGCAT 2374  
 Oy 755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775  
 Db 2375 GGGCTATGAGCGGGTCCAGAGCGCTCTTCCGAGCGCTGATGTGACCTCATATACCT 2434  
 Oy 775 uIleThrSerAspTyrGlnLeuLeuSerAspProThrProGlyLleProSerProArgAs 795  
 Db 2435 CATCTCTTCAACTATGAGCTCTCTCAGACCCACACCTGTGTGCTGTGCACCTGTGTA 2494  
 Oy 795 pGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnG 815  
 Db 2495 TGGGCTGTGG--AATGTTGCCAGCTCTATGCTGTCMAAGACCCCAAGCATCTTCAGAGA 2551  
 Oy 815 uArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835  
 Db 2552 GAGACACCTCAAGATACATCTCAGACGCTGGGCAAGGGCAACTTTGGCAGCTGTGAGCTGTG 2611  
 Oy 835 sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnH 855  
 Db 2612 CCGCTATGACCCGCTACCCCAATACAGTGGCTGTGTGGCGGTGAACACGTGTGAGCA 2671  
 Oy 855 sSerValProAspGlnGlnArgAspPheGlnArgGlnGlnIleGlnIleLeuLysAlaLeuH 875  
 Db 2672 CACGGGCGCAGACAGAGAGGAGTTCAGCGGAGATTCAGATCTTCAAGACACTGTGA 2731  
 Oy 875 sSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895  
 Db 2732 CAGTGAATTTATGTGTAAGTATGTGTGTGACACTTAGGCGCGGCGGCGCAGAGCTGGG 2791  
 Oy 895 gLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgG 915  
 Db 2792 GTGTGATATGAGTACTGTGCGGAGCGGTGTGCGGCGGACTTCTGTGACGCGGACCGCGC 2851

QY	915	Y--L	LeuHisThrAspArgLeuLeuLeuLeuPheLeuIleATPrGInIleCysLysGlyMetGlu	934
Db	2852	GGCGCTGCATGCCCGCGCCCTCTTCTCTATCTCTCCAGATGTGGAAAGGGCATGGAGTA	2911	
QY	934	rLeuGlyAlaIArgArgCysValHisAArgAPLeuAlaAlaArgAsnIleLeuValGlu	954	
Db	2912	CTGAGGCTCCCGCGCGCTGCTGGCACCGCGACCTGGCCCGCGAACAATCTCTGGAGAG	2971	
QY	954	rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLys	974	
Db	2972	CGAGGCACACGCTCAAGATCGTGACTTCGGCTTAGCTTAAGCTACGCTGCGCGTTAGAAAGA	3031	
QY	974	pTyTyTyValValArgGluProGlyGlnSerProIlePheTrpTyAlaProGluSerIle	994	
Db	3032	CTACTACCTGGTCCGCGAGCGACGCGACAGCCCATTTTCTGATGATCCCGCGAATGCTT	3091	
QY	994	UserAspAsnIlePheSerArgInSerAspValTrpSerPheGlyValValLeuTyrl	1014	
Db	3092	CTCGGACAAACATCTTCTCTGCCAGTACAGAGCTGTGGAGCTTCGGGCTGCTGTACGA	3151	
QY	1014	uLeuPheThrTyrcysAspIlysserCysSerProSerAlaGluPheLeuArgMetGle	1034	
Db	3152	GCTCTTCACCTACCTCGACAAAGCTGACGCCCTCGGCGCAATTCCTCGGAGATGGG	3211	
QY	1034	YProGluArgGluGlyProProLeuLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArg	1054	
Db	3212	ATGTGAGCGGGATGTCCTCCGCCCTCTGCCGCTTGGACATGCTGGAGGAGGCGCAG	3271	
QY	1054	gLeuProProProProIlnProIlnCysProIlnGluValGlnIleuMetGlnLeuCysTrpAl	1074	
Db	3272	GTGTGCGCGGCTCTGCTGCTGCTGCTGCTGAGGTTTACAGACTCATGAAGCTGTGCTGGCC	3331	
QY	1074	aProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTr	1094	
Db	3332	CCCTAGGACCAAGACCGGCGCATTCATGACGCCCTGGGCGCCCGACGCTGACATGCTGTG	3391	
QY	1094	pArgGlyArgProGly	1099	
Db	3392	GAGCGGAGCGCGGGG	3407	
RESULT 7				
LOCUS	HSU31601		3620 bp	mRNA
DEFINITION	Human tyrosine protein kinase (Jak3b) splice variant mRNA, complete cds.			
ACCESSION	U31601			
VERSION	U31601.1			
KEYWORDS	GI:1039418			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens.			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 3620)			
JOURNAL	Lai, K.S., Jin, Y., Graham, D. K., Witthuhn, B. A., Ihle, J. N. and			
MEDLINE	Liu, E. T.			
PUBMED	A kinase-deficient splice variant of the human JAK3 is expressed in			
REFERENCE	hematopoietic and epithelial cancer cells			
AUTHORS	J Biol Chem. 270 (42), 25028-25036 (1995)			
TITLE	2 (sites)			
JOURNAL	Witthuhn, B. A., Silvenoinen, O., Mura, O., Lai, K. S., Cwik, C.,			
MEDLINE	Liu, E. T. and Ihle, J. N.			
PUBMED	Involvement of the Jak-3 Janus kinase in signalling by interleukins			
REFERENCE	2 and 4 in lymphoid and myeloid cells			
AUTHORS	Nature 370 (6485), 153-157 (1994)			
TITLE	3 (bases 1 to 3620)			
JOURNAL	Liu, E. T.			
MEDLINE	Direct Submission			
PUBMED	Submitted (14-JUL-1995) Edison T. Liu, Linberger Comprehensive			
REFERENCE	Cancer Center, University of North Carolina, Chapel Hill, NC			
AUTHORS				
TITLE				
JOURNAL				

COMMENT	27599-7295, USA
Wittuhnh, B.A. Nature 370, 153-157, 1994.	
FEATURES	
source	Location/Qualifiers
	1..3620
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="19"
	/map="19p12"
	/clone="SKBR-3"
	/cell_type="SKBR-3 breast cancer cells"
gene	1..3620
	/gene="jak3b"
CDS	96..3380
	/gene="jak3b"
	/function="tyrosine protein kinase"
	/codon_start=1
	/product="JAK3b"
	/protein_id="AAC50226.1"
	/db_xref="GI:1039419"
	/translation="MAP3PSEPTLIPORSCSILSTFAGALHVLPAAGPPORLSFS FGHLEADLCVQAKASGILPVYHSJLPALATEDLSQWPSHSIFSEVDASTOYLRI RFEYPMWGLEKCHRFGLRKRDASATLIDLVLEHLRQHRSDVSGRSLPGVSLTKBOG ECISLIVLDIARMAREQAPRQGBELKTVSFKACLPRLDLOGLSVTRRRIRRYR RALRNVACQADSHSLMAKYIMDLERIDPAGAEETPVHGLPAGLGGHDGGLLVAD GGIIMVQEOEVLQPCDPEIIVDISCKAPRAGVEHNLVYTRIDNOIILAEPRG LPELISVALVADGFRITDTSOHFCEVAPRFLLEVAQOQGPITLDPAIKRLG GSRGVSIVLRSRDPGDSFLITVQVONLPADPKYKIRSPGTFLLVGLSPHSL RELATITWDSGLHYDVAVITLISCCITPRCKRSNLIYVQGHSEVGRKTEVILKVDANKNM SOMTEKIPADSDSEHNLGHSGFTKTYRCCREHYNGEARKTEVILKVDANKNM ESPELEASLMSQVSRYHLVLLHGVCMAGDSTVOEYVHLCAIDYILRCHVLPASMK LOVVQLAVLANYLEDGKLPNGHVSARKVILLARGADSPFKLSDPGSPAVLSLE MLVRIPIVPAVECLREAOITLSLEADKMGFGATWVEVSQVTPILSDPKKLOFTED ROOLPAKTEELALLIOOCMAVEVQRPSPRAVIRDNLSISDYELSDPTGALVQ RDGIMGNAQIYACODPTIFEEERHLKYSIQLGKNGFSGVEICRDPLDNDTPALVQ LQHGSPQOQRDFOREIOLIKALHSDFIYVKRYGSGROSILRYMVEYLSGLRDL QRHARLADSLHSLSSQICKMGWEIVDSRCVHRDLAARNILVSEAHVILADEGLAK LLPLDKRTYVIREPGQSFITMYAPESLSDNIFESQSPWMSGVLYELFETVCDKCSPL SAEELIRMGCEBDVPALCRLELLEESQRLEPAPCAFEVSAAGLASVDSVDMAVS GKPAGA"
BASE COUNT	670 a 1152 c 1076 g 722 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.01e-273
Score:	4597.00
Percent Similarity:	86.80%
Best Local Similarity:	80.92%
Query Match:	78.45%
DB:	9
	Length: 3620
	Matches: 895
	Conservative: 65
	Mismatches: 123
	Indels: 14
	Gaps: 7
US-09-397-967-16 (1-1099) x HS031601 (1-3620)	
QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20	
Db 96 ATGGCAGCTCCAGTGAAGAAGAGCGCCCTCGATCCCTCACAGCTTCAAGCCCTTGTGCC 155	
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40	
Db 156 ACGGAGCGCTGGTCCCTGCTGCTGTGCTGTGCCGCTCGGGGCCCGGGCCCCCCCCAGCGCC 215	
QY 41 LeuSerPheSerPheGlyAspTyrIleuAlaGluAspLeuCysValAlaArgAlaIalaValAla 60	
Db 216 CTATCTTTCTCTTTGGGAGCACCTTGGCGAGAGACCTGTGCTGACAGGCTGCCAAGGCC 275	
QY 61 CysGlyIleLeuProValAlaTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80	
Db 276 AGCGGCTCTGCTGTGTACCACTCCCTTTGTCTGTGGCCACGGAGAGCACTGTCTGCC 335	
QY 81 TrpPheProSerSerHisIlePheCysIleGluAspValAspThrGluValLeuValTyr 100	
Db 336 TGGTTCCTCCCGAGCCACATCTTCTCGGAGAGATCCAGCACCCAAAGCTCTGTCTAC 395	
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluIleuThrCysHisArgPheGlyLeu 120	

Db 396 AGAATTCCTTTACTTCCCAATTTGGTTGGCTGGAGAAAGTCCACCCCTTCGGGGCTA 455  
Oy 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
Db 456 CGCAAGGATTTGGCCAGGCTATCTTGGACCTGCGACCTGGAGCACCTCTTTGGCCAG 515  
Oy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
Db 516 CACCGCAGTGAACCTGTGAGTGGGGCCCTCCCTGGGGCTTCAAGTCAAGAGAGAGGT 575  
Oy 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
Db 576 GAGTGTCTCAAGCTGGCCGTGTGGACCTGGCCCGGATGGCCGAGAGAGCCGAGCCGCG 635  
Oy 181 ProGlyLeuLeuLeuLysThrValSerTyrLysAlaCysLeuProSerLeuArgAsp 200  
Db 636 CCGGAGAGCTGCTGAAAGACTGTACACTACAGGCTGCTCCCTCCCAAGCTGGCGGAC 695  
Oy 201 ValIleGlnGlyLysAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220  
Db 696 CTGATCCAGGGCCTGAGCTTCTGACGCGAGGCGTATTCGGAGAGCGGTGGCGAGAGCC 755  
Oy 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240  
Db 756 CTGCG-CCCGCTGGCCCTGCCAGGAGACCGGCACTGCTCATGGCCAGTACATCAT 814  
Oy 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGly 260  
Db 815 GGCACCTGGAGCGGCTGATCCAGCCGGGCGCCGAGACCTTCCAGTGGGCTTCCCTGG 874  
Oy 260 yAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277  
Db 875 GCGCCCTGGTGGCCAGACGGGCTGGGCTGCTCGCGCTGGCTGTGACGGCCGACATCC 934  
Oy 277 OTPrSerSerAsnAsp--GluLeuPheGlnThrPheCysAspPheProGluIleLeuAla 296  
Db 935 CTGGACCCAGGAGAAACAGAGAGTCTCCAGCCCTTCTGGACATTTCCAGAAATCGTAGA 994  
Oy 296 rValSerIleAsnGlnAlaProArgValGlyProAlaGlyLysHisArgLeuValThrVa 316  
Db 995 CATTTAGCATCAAGCAGGCGCCGCGGTGGCCGCGCGGAGACCGGCTGGTCACTGT 1054  
Oy 316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336  
Db 1055 TACCAGGACAGACAACAGATTTTAAAGCGGAGTCCCAAGGCTGCCCGAGGCTGTGC 1114  
Oy 336 rPheValIleLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356  
Db 1115 GTTGTGGCGCTCGTGGACGGCTACTTCCGGCTGACCACCGACTCCACAGACTTCTTCTG 1174  
Oy 356 sLysGluValAlaProProArgLeuLeuGluGluGluAlaAspValCysHisGlyProI 376  
Db 1175 CAAGGAGGTGGACCGCGGAGGCTGTGAGAGAAAGTGGCGAGACAGGACGACGCGCCCAT 1234  
Oy 376 eThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrI 396  
Db 1235 CACTCTGACTTGTCCATCAACAAGCTCAAGACTGGGGGCTCAACGCTCGGCTCATGT 1294  
Oy 396 eLeuAlaArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416  
Db 1295 TCTCCCGCGAGCCCGCAGAGACTTTGACAGCTTCTCTCACTGCTGTGTCAGAAACC 1354  
Oy 416 oLeuGlyProAspTyrLysGlyCysLeuIleArgIleAspProSerGlyAlaPheSerIe 436  
Db 1355 CTTTGGTCTGATATTAAAGGCTGCTCATCCGCGCAGCCGACAGGAAAGCTCTTCTT 1414  
Oy 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTTrpAsnSe 456  
Db 1415 GGTGGGCTCAGGCCAGCCACAGCAAGTCTTGGAGAGCTCTGGCAACCTCTGGGATGG 1474  
Oy 456 rGlyLeuArgValAspGlyAlaAlaLeuLeuTyrLeuThrSerCysValaProArgProLy 476

Db 1475 GGGCTGCACCGATATGGGGTGGCAGTACACCTTCTCTGCTATTCGCCAGACCCAA 1534  
Oy 476 sGlyLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCy 496  
Db 1535 AGAAAGATCCAAACCTGATCGGTGCCAGAGAGGTCCACAGCCCAACATCATCTTGGT 1594  
Oy 496 sSerPro--SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515  
Db 1595 TCAGCCCAATCCCAATACAGCTGAGTCAAGATACATTTCCACAAGATCCCTGACAG 1654  
Oy 515 rLeuGluThrPheIleGluAsnLeuGluHisGlySerPheThrLysIlePheArgGlySer 535  
Db 1655 CTTGAGAGTGGCATGAGAACCTGGGCGCATGGGCTCTTACCACAAGATTACCGGGGCTGTG 1714  
Oy 535 gArgGluValAlaAspGlyGlnThrHisAspSerGluValLeuLeuLysValMetAspSe 555  
Db 1715 CCATAGAGGTGTGATGGAGAGGCGCCGAAACAGAGAGTCTCTGTAAGGTCTCATGGATGC 1774  
Oy 555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575  
Db 1775 CAACACAAAGAACTGATGAGATCATTTCTGGAAGCAGCAGCTTATATACCCAGAGTTC 1834  
Oy 575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValG 595  
Db 1835 GTACCGGCAATCTGTGCTGTCCACGGCGGTGTGCATGTGAGAGACAGCAGCATGGTGA 1894  
Oy 595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615  
Db 1895 GGAATTTGTACACCTGGGGGCCATTAACATATATCTGCCAAACGAGGCGCACCTGGTCC 1954  
Oy 615 rAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAs 635  
Db 1955 AGCCACTGGAACTGACAGTGTGCTCAACAGCTGGCTACGCCCTCAACTATCTGAGAGA 2014  
Oy 635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGly 655  
Db 2015 CAAGAGCTGCCCATAGGCAATGTCTGTCCCGGAAGTGTCTCGCTCGGAGGCGGC 2074  
Oy 655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675  
Db 2075 TGAATGGAGCCCGCTTATCATCACTGATGACCTTGGGCTGACGCCGCTGTGTATAG 2134  
Oy 675 rLeuGluMetLeuThrAspArgIleProThrValAlaProGlyCysLeuGlnGluAlaG 695  
Db 2135 CCTGGAGATGCTCACAGCAGAGATCCCTGGGTGGCCCGCAGTGTCTCCGGAGGCGCA 2194  
Oy 695 rThrLeuCysLeuGluAlaAspLysTPrGlyPheGlyAlaThrThrTrpGluValPheG 715  
Db 2195 GACACTTGAAGCTTGAAGCTGACAGTGGGCTTGGCGCCACGGTCTGGGAAGTGTTAG 2254  
Oy 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735  
Db 2255 TGGCCTCAACATGCCCATCACTGCTGTGATCTGTGTAAAGAACTCCAAATTTATGAGA 2314  
Oy 735 pGlnGlyLysLeuProAlaLeuLysTPrThrGluLeuAlaGlyLeuIleThrGlnCysMe 755  
Db 2315 CCGGAGAGAGCTGGCGGCGCCCAAGTGAACAGACTGGCCCTGTGTAAACAGTGCAT 2374  
Oy 755 tAlaTyrSerProGlyLysArgProSerPheArgAlaIleLeuAlaArgAspLeuAsnGlyLe 775  
Db 2375 GGCCTATGAGCGGCTCCAGAGGCTCTCTTCCGAGCGCTATTCGTGACTCAATAGACT 2434  
Oy 775 uIleThrSerAspTyrGlyLeuLeuSerAspProThrProGlyIleProSerProArgAs 795  
Db 2435 CATCTCTTCAAGATATAGCTCTTCAGACCCACACCTGGTGGCCCTGGCACCTCTGTA 2494  
Oy 795 pGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluG 815  
Db 2495 TGGGCTGTGG---AAATGATCCCAAGCTCTAATGCTCCCAAGACCCACAGATCTTCGAGA 2551  
Oy 815 uArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835  
Db 2552 GAGACACTCAAGTATCATCTCAAGCTGGGCAAGGCAACTTTGGCAGGCTGGAGCTGTG 2611

QY 835 sargTyrAspProLeuGlnLysAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnH 855  
 Db 2612 CCGGTATGACCCCGGTAGCGGACATACAGGTGGCCGTGGTGGCGGTGAACAGCTGCAGCA 2671  
 QY 855 sSerValProAspGlnGlnArgAspPheGlnArgGlnLysGlnLeuValLeuH 875  
 Db 2672 CAGCGGACACACGAGAGAGGACTTTCACGGGAGATTCAGATCTCCTCAACAGCAGTGA 2731  
 QY 875 sSerAspPheLeuValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895  
 Db 2732 CAGTATTTCATTGACAGATATCTGGTGTCAGCTATGCGCGGCGGCGGAGACCTGCGC 2791  
 QY 895 GlnValMetGlnTyrLysLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgG 915  
 Db 2792 GCTGTCTATGAGATGATCTGCGGAGGCTGCTTGGCGGACTTCTGCGAGCGGCGGCGC 2851  
 QY 915 Y---LeuHisThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGly 934  
 Db 2852 GGGCTCGATGCCAGCCCGCTCTCTATCTCGGAGATCTGCAAGGCGCATGAGTA 2911  
 QY 934 GlnValAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValG 954  
 Db 2912 CTTGGGCTCCCGCGCTGCTGCGACCGGAGCTGGCGCGGCGGCGGCGGCGGCGGCGGCGG 2971  
 QY 954 rgluaLHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974  
 Db 2972 CGAGGACACGTCAGATGCTGACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3031  
 QY 974 pTyTyrValValArgGlnProGlyGlnSerProIlePheTrpTyrAlaProGlySer 994  
 Db 3032 CTACTAGGTGGTCCGCGAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3091  
 QY 994 uSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrG 1014  
 Db 3092 CTGCGAACATCTCTCTGCGCAGTCAGACGCTGAGCTGCGGCTGCTGCTGCTGCTGCTG 3151  
 QY 1014 uLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMet 1034  
 Db 3152 GCTCTTCACTACTGCGCAAAAGCTGACGCCCTCGCGGAGATTCTCGGATGATGAG 3211  
 QY 1034 YProGluArgGlnGlyProProLeuGlnCysArgLeuGlnLeuAlaGluGlyArg 1054  
 Db 3212 ATGAGACGGGAGTCCCGCCCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3271  
 QY 1054 GlnProProProProThrArgProThrGluValGlnLeuMetGlnLeuCysTrpAl 1074  
 Db 3272 GCTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3305  
 QY 1074 aProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1094  
 Db 3306 -----ACGCCCGGAGGCGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3346  
 QY 1094 pArgGlyArgProGly 1099  
 Db 3347 GCGCGGCGTCTGCGG 3362  
 RESULT 8  
 AC073700 218208 bp DNA linear HTG 29-JUN-2000  
 LOCUS AC073700 Mus musculus clone RP23-162012, WORKING DRAFT SEQUENCE, 16  
 DEFINITION Mus musculus clone RP23-162012, WORKING DRAFT SEQUENCE, 16  
 unorderd pieces.  
 ACCESSION AC073700  
 VERSION AC073700.1 GI:8810317  
 KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 218208)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Sequencing of Mouse  
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 218208)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 COMMENT  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: <http://www.jgi.doe.gov>  
 Project Information  
 Project Name: 1799428  
 Center clone name: RPI-23\_162012  
 -----  
 Summary Statistics  
 Consensus quality: 204328 bases at least Q40  
 Consensus quality: 210800 bases at least Q30  
 Consensus quality: 212488 bases at least Q20  
 Estimated insert size: 22200; agarose-1p estimation  
 Estimated insert size: 216708; sum-of-contigs estimation  
 Quality coverage: 9.37 in Q20 bases; sum-of-contigs estimation  
 Quality coverage: 9.6 in Q20 bases; sum-of-contigs estimation  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 16 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1  
 1093 1092: contig of 1092 bp in length  
 1193 1192: gap of unknown length  
 1193 2565: contig of 1373 bp in length  
 2566 2665: gap of unknown length  
 2666 3711: contig of 1046 bp in length  
 3712 3811: gap of unknown length  
 3812 5555: contig of 1744 bp in length  
 5556 5656: gap of unknown length  
 5657 9014: contig of 3359 bp in length  
 9015 9114: gap of unknown length  
 9115 12524: contig of 3410 bp in length  
 12525 12625: gap of unknown length  
 12626 18315: contig of 5691 bp in length  
 18316 18415: gap of unknown length  
 18416 25299: contig of 6884 bp in length  
 25300 25399: gap of unknown length  
 25400 34958: contig of 9459 bp in length  
 34959 34859: gap of unknown length  
 34860 48111: contig of 13153 bp in length  
 48112 48212: gap of unknown length  
 48213 59675: contig of 11464 bp in length  
 59676 59775: gap of unknown length  
 59776 79239: contig of 19464 bp in length  
 79240 79339: gap of unknown length  
 79340 102759: contig of 23420 bp in length  
 102760 102859: gap of unknown length  
 102860 129973: contig of 27114 bp in length  
 129974 130073: gap of unknown length  
 130074 164188: contig of 34115 bp in length  
 164189 164289: gap of unknown length  
 164290 218208: contig of 53920 bp in length.  
 FEATURES  
 source 1..218208  
 1..218208  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="RP23-162012"  
 /clone\_lib="RPI mouse BAC library 23"  
 BASE COUNT 51066 a 56856 c 56653 g 52132 t 1501 others  
 Alignment Scores:  
 Pred. No.: 5,71e-240 Length: 218208  
 Score: 4091.50 Matches: 1072

Percent Similarity: 32.75%  
 Best Local Similarity: 32.69%  
 Query Match: 69.82%  
 DB: 2  
 Conservative: 2  
 Mismatches: 24  
 Indels: 2188  
 Gaps: 24

US-09-397-967-16 (1-1099) x AC073700 (1-218208)

QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSer 20  
 Db 110153 ATGGACCTCCAACTGAGAGACACCTCTGATCCCTCAGCGCTCTGACCTCTCATCC 110212  
 QY 21 SerIleuAlaGlyAlaLeuHisValLeuLeuProProArgIleProGlyProGlnArg 40  
 Db 110213 TCACAGGCGAGAGCCCTGCATGTGCTCTCTCCCGGGGAGCCTGGGCTCCCGACCA 110272  
 QY 41 LeuSerPheSerPheGlyAspIleuAlaGluAspLeuCysValArgAlaIleuAsp 60  
 Db 110273 TTGTCATCTCTTTTGGGGACTACTTGCGCTGAGGATTTATGTGTGCGAGCTGCCAAGGCC 110332  
 QY 61 CysGlyIle----- 63  
 Db 110333 TGTGTGA-GTGGTCCCTAGATGACCATGCATGTATAGCAACATCTATTGAGAAGG 110391  
 QY 63 ----- 63  
 Db 110392 GGATGCTGGGCTTTTGAAGGGGTGTGAATTTGGTAAGGAGAGATCAGATCAGCCGG 110451  
 QY 64 -----LeuProValIleHis 68  
 Db 110452 TGTAGTTGACTTGGTGAACTGATGCCCCATCCCTTCAGGACATCTGCTGTTATCAT 110511  
 QY 69 SerLeuPheAlaLeuAlaThrGluAspPheSerCysIleProPheProSerHisIlePhe 88  
 Db 110512 TCCTTTTGGCTGTGGCCACTGAGGACTTCTGTGCTGTTTCCCGCAAGCACATCTTC 110571  
 QY 89 CysIleGluAspValAspIleThrGluValIleuValIleuValIleuValIleuVal 103  
 Db 110572 TGCATAGAGAGAGTGAGACCTCAAGCTTGTCTACAGGCTTACGAGTGAAGACTGAACCC 110631  
 QY 103 ----- 103  
 Db 110632 TTGATTTATCTTCAAGGATGATTTTCTATTCAGTCTATAGTAGAGGCTTTGAACAGA 110691  
 QY 103 ----- 103  
 Db 110692 GAGGTTGGGCTGTGCTGTAGGCCACACTGCATTCAGTACCCAACTTGGGCTGTGG 110751  
 QY 104 -----PheTyrPheProAspTyr 109  
 Db 110752 AAGCCAGAGCTGAGTGGGCCAACCCCATTTCTGAAACCCACAGCTTTTATTTCCCTGACTGG 110811  
 QY 110 PheGlyLeuGluThrCysHisArgPheGlyLeuArgGlyAspLeuThrSerAlaIleLeu 129  
 Db 110812 TTTGGGCTGGACATGTCTACCGCTTTGGCTGGCGCAAGATTTGACACAGTGCATCTCT 110871  
 QY 130 AspLeuHisValLeuGluHisIleuPheAla----- 139  
 Db 110872 GACTTACACGTTTGTAGAACATCTCTTGTCTCAGGTGGTATTCTTCCGTGGGGAAGCTC 110931  
 QY 139 ----- 139  
 Db 110932 AAGCCAGGATGGGGAGCCCAAGGTAAACAATCCAGACCCAGACCGAAGCTCACT 110991  
 QY 140 -----GlnHisArgSerAspLeuValSerIleArgLeuProVal 152  
 Db 110992 AAGCACTCTCTGTGGCTGACCCAGACACCCAGTACCTGTGGTGGGCGCTCCCGGT 111051  
 QY 152 IGLYLeuSerMetIleGluGluIleuPheLeuSerLeuAlaValLeuAspLeuAlaI 172  
 Db 111052 GGGGCTTAGCTGAAAGAGCAGGAGAGATTCTGAGCCTGCGCTGTGACTTGGGCCA 111111  
 QY 172 MetAlaArgGluGlnIleGlnArgProGlyGluLeuLeuIleuVal----- 188  
 |||||||

Db 111112 GATGCTCGTGAAGAGCCAGCGCCAGAGAGAGCTGTGAAGAGGTACAGTGAAGGCC 111171  
 QY 188 ----- 188  
 Db 111172 TCGGGCCACCCTTGTGTGCTTCTTTGGCTGGAGATAGGCGCTATGCTTTCCCTT 111231  
 QY 188 ----- 188  
 Db 111232 ACTGAAACATACCTCTTCCCACTTGTGAGCACCCTCTCCGTAGCCTGCCCTG 111291  
 QY 188 ----- 188  
 Db 111292 CCGAAGGTCCCTATGTCTGTCTGTCTCTGGGCCCCCTCACCACTTAAGAC 111351  
 QY 189 -----SerTyrIleAspIleuProProSerLeuArgAspVal 202  
 Db 111352 TCGGGCTGCCCCCCCCCCCAGATTACAAAGCTGTCTGCCCCAGACCTGGCGCATGTGA 111411  
 QY 202 IeGlnGlyGlnAspPheValIleuArgArgIleuArgArgIleuValIleuAlaLeu 222  
 Db 111412 TCCAGGGCCAGAACTTCTGTACAGCAGCGCATCCGAGAGCGGTGTGGCGCTGC 111471  
 QY 222 euProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaIleuAsp 241  
 Db 111472 G-CCGTGTGTGCTGCTGCGAGCGGACCGCTACGCGCTCATGGCCAAATATTGTGAC 111530  
 QY 242 LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValIleuProGlyVal 261  
 Db 111531 CTGGAGCGGCTAACATCCAGCGGCGCACACGAGACCTTCGTGTGGGGCTCCCGGGCGCC 111590  
 QY 262 GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerAsn 281  
 Db 111591 CAGAGAGGCGGGGCTTCTGTGTGTGGGGGACAAACGATCTCTGAGCTCGGG 111650  
 QY 282 AspIle----- 283  
 Db 111651 GACCA-GGAGGTGTGGGGCTGGGCTTGAAGAGCTGTCCCGGGAGCGGTGGCGGT 111709  
 QY 283 ----- 283  
 Db 111710 TAGTAAAGAACCCCTCCAGAACATCAGAAAGTGTGTGCGGGCGGAGTGTGA 111769  
 QY 283 ----- 283  
 Db 111770 CTAGGGGCGAGGGGGGGGCTCAATTAAAGTGTGGTCCGGGATGAGCTACCCAGGA 111829  
 QY 283 ----- 283  
 Db 111830 CTGGAGCTGGGGCGTGAAGAGGTGAGACTAGATGAGGGGGCGGAGCGAGAGAG 111889  
 QY 283 ----- 283  
 Db 111890 GCGATAGTATGATTTGGATCGGTGGGTGTGACCCGACAGGGCTGTACGTCCCTAAC 111949  
 QY 283 ----- 283  
 Db 111950 CTGTCGCCCGACCCATCCACCCCTCCCGCGTAAAGTGGAGTGGGAAGAGAGAG 112009  
 QY 283 ----- 283  
 Db 112010 AAGGAGAGCTGTGTGTGAGGTCTGTAATCTGTGCTTCTGCTGTCTCTTCT 112069  
 QY 283 ----- 283  
 Db 112070 CTTTATTTTATGTTTGGGAGACAGGTTTCTGTATAGCTGTGGCTGTCTGGA 112129  
 QY 283 ----- 283  
 Db 112130 ACTCACTTTGTAGACAGGCTGGCTGTGAACTCAGAAATCCGCTGTGCTCTGCA 112189  
 QY 283 ----- 283  
 Db 112190 GTGCTGTATTAAGCGGTGGCGCACTACACCAAGCTCTGCTGTCTTCTCTCAG 112249

Oy 284 LeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaPro 303  
 Db 112250 CTTTTCAGACCTTCTGTGACTTCCGAAAAATCGTGGATGTGCAATCAAGAGGCCCA 112309  
 Oy 304 ArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIle 323  
 Db 112310 CGTGTGGCTCCGGCAGGGAACCGGCTGTCCACCGTCCAGGATGAGCGGCACATC 112369  
 Oy 324 Leu----- 324  
 Db 112370 CTGGTGGCTCTGGCACCCTTATTCTGTCTGATCTCTGGAAAGGGGTCCGTT 112429  
 Oy 325 -----GluAlaGluPheProGlyLeuProGluAlaLeu 336  
 Db 112430 GTTCTGAGAGTTCTCTGGCGCAGGAAGCGAGTTTCCGGGGCTGCTGAGCGCTGTC 112489  
 Oy 336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCys 356  
 Db 112490 CTTCGTGCTCTGTGTGATGGGTACTTCCGCTGATCTGCGACTCCAGGCATTATTCTG 112549  
 Oy 356 slySGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIle 376  
 Db 112550 CAGAGAGGTGGCGCGCGCGGCTGTGTGAGAGAGAGCGGAGCTGTGCATGAGCCAT 112609  
 Oy 376 eThr----- 377  
 Db 112610 CACGTTAGAGCTAGATTGGGTGACAGTGGGAGGGGCTGGCAGCAGAGCTGTCTGG 112669  
 Oy 378 -----LeuAspPheAlaIleHisIleLeuValAla 388  
 Db 112670 GTTGGACCTGGCTGCTGCTTGGCTTGGCTTGGATTGACTTGGCATCCAAAGCTGAAGCGCG 112729  
 Oy 388 IagIleSerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspPheIle 408  
 Db 112730 CTGGCTCCCTCCAGCAGCACTACATTCTCCGCCGACGCCAGAGACTATGACAGCTTTC 112789  
 Oy 408 euleuThrAlaCysVal----- 413  
 Db 112790 TTTCTTACCGCTGCGT-CCAGSTTGTGTTACTGTGTGTCAGAGATGGCTTCTGTGAGCA 112848  
 Oy 413 ----- 413  
 Db 112849 AGTGACACTGAGCTGGGACAGAAAGAGCGTTCTATTCTTCTTGAAGACACTGGCC 112908  
 Oy 413 ----- 413  
 Db 112909 TGTAACTCAGATTTAAGAGACTGAGCAGATCAAGTTCAGAGCCTGCTGAGCTAT 112968  
 Oy 413 ----- 413  
 Db 112969 ATTAGAAATTAATGTTTCTCATGAGGAAAGGTTTTTCCACCAAGCCTGAGGGTGT 113028  
 Oy 413 ----- 413  
 Db 113029 CTTTCACTCAGCATTCACAGCAGTGAAGACAGACACTGACATTAATCGGCTCG 113088  
 Oy 413 ----- 413  
 Db 113089 ACCTCCACAGTGCAGGTGACTTGCATCTACACAGAAACACATTAAGTAATATT 113148  
 Oy 413 ----- 413  
 Db 113149 TAAGAACTCGAAGAGCTCAGAAACAGTCTCAGCATGGCGGGGGGGGAGAG 113208  
 Oy 413 ----- 413  
 Db 113209 CATCTTAGAGTTGGAGACAGAACTGAGCATCACTTCCAGAAAGAGAGGGGT 113268  
 Oy 414 -----GlnThrProLeuGlyP 419  
 Db 113269 GGGTAGCTGTCTCCACATCTGGGTCAGTGAAGTCTCTTGTGAGAGACTCTCTGTGGCC 113328

Oy 419 roAspTyrIleCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyL 439  
 Db 113329 CCGACTACAGAGGCTGCTCATCTCCGACAGACCCAGGCGGGCTTCTCCCTGGTTGGCC 113388  
 Oy 439 euSerGlnProHisArgSerLeuArgGluLeuAlaIleCysTrpAsnSerIleLeu 459  
 Db 113389 TCAGCAGCCCAACAGAAAGCTGCGGAGCTGTTCAGCTGTGGAATCTGGCTGC 113448  
 Oy 459 rValAspGlyAlaIleLeuTyrIleThrSerCysAlaProArgProIle 476  
 Db 113449 GAGTAGACGTCTGCTCCAGAACTAATCTCTGCTGCTCCAGACCAAGGGTGAAT 113508  
 Oy 476 ----- 476  
 Db 113509 CAGTTCTCCCTAACACAGCCCGCAGCTCTGGGTGTAGCAGAAACCAAGACCATGG 113568  
 Oy 476 ----- 476  
 Db 113569 CGATTAGACAGAGAACACGACTTTTGGCACAGTCTTCAGACATCAAGTCAATTTT 113628  
 Oy 476 ----- 476  
 Db 113629 CAGTTGGCGCGTCTTCCCATGGGTGATGCCACTAAGCGAAACCTGCTGGACA 113688  
 Oy 476 ----- 476  
 Db 113689 CATCTCCCTCTGTGTGTCTGTGCTTACAAAGAACCCAGCGGCTGGAGGAGACCCAGT 113748  
 Oy 476 ----- 476  
 Db 113749 ACGTTCAATGTTCAANTCAGCCCAAGTCAAGATCTCAGAGCGACTTCCGGGGCGCC 113808  
 Oy 476 ----- 476  
 Db 113809 CCAAGCTAAGTAGAGAGTGAATCTTGGAAACCTTGTGCTGACTGCTTTCTT 113868  
 Oy 477 -----GlyuSerAsnLeuIleValAlaArgArgGlyCysAsnProAlaProAlaP 494  
 Db 113869 GCCATCCAGAAAGTCCAAATTTATGCTGTGGGAAGGGGTGACACCCGGCGCTGCC 113928  
 Oy 494 roGlyCysSerProSerCysAlaLeuThrGlnLeuSerPheHisThrIleProThr 514  
 Db 113929 CTGGCTGCTCCCGCTCTGCTGTGCTGACACAGCTGACCTTCCACAAATTCACAGG 113988  
 Oy 514 sPserLeuGluTrp----- 518  
 Db 113989 ACAGCTGAGTGGGTAGAGGTCCCAAGGTGAAGCCAGGCGCGCTTTTGGAGG 114048  
 Oy 519 -----HisGluAsnLeu 522  
 Db 114049 ACGTGAAGCCACTTCAACTGTCTCGCTGGGTGTCTCCCGCAGCAGGAACCTG 114108  
 Oy 523 GlyHisGlySerPheThrIlePheArgGlySerArgArgGluValAlaAspGlyGlu 542  
 Db 114109 GGTACGGTCTTTTACCAATCTTCGTGGCCGACGGCGGAGGTCTGTGATGTGTGAG 114168  
 Oy 543 ThrHisAspSerGluValLeuLeuValMetAspSerArgHisArgAsnCysMetGlu 562  
 Db 114169 ACACATGACTGGAAAGTCTCTGAGAGTATGAGATCCAGACATCGGAATCTGATGAG 114228  
 Oy 562 ----- 562  
 Db 114229 GTTCGACGGGACAGAGAGGCTTGCAGAACTGCATGGAATCTGGACCTGGCAGAGG 114288  
 Oy 562 ----- 562  
 Db 114289 GTTACTGGAACCTAGAACTTAAGACACGACTGGGGGACATAGTGGGACTCCATCAACC 114348  
 Oy 562 ----- 562  
 Db 114349 AAAGTAGATTAGCATGCTGTGATGTTGTAACAAACAGCATGAGAGGCTGGGAGG 114408  
 Oy 562 ----- 562

Db 11548	CGGGGTGTACAGATGGAACCTGGTGAAGAAGGGGGTGTCTTCAATGCGCGATGTGGG	11554
Qy 632	-----	632
Db 11554	CTTTAACTGTCCTTAAAGTCGTGGAGAGTACAGAAAGGGGTGTGGTCTTAAAGATGGG	11560
Qy 632	-----	632
Db 11560	TATGTCAGTGTAAAGGGAGGTTAAGGTTACAGACATGACTTACAGGGCTACAGTGGT	11566
Qy 632	-----	632
Db 11566	TGCGAAAGCCTGTTGTGTGTAGGCAGCATAGATAGGCTTTAGAGATGACAGAGATT	11572
Qy 633	-----	633
Db 11572	GCTCAACAACATACTTCCCTGTCCTTCCTTACCACTTAAAGAGGACAAAGGCTTCCCA	11578
Qy 640	-----	640
Db 11578	CGGCAAGCTCTACACAGGAGGTCCTGCTCGTGTGAGGGGGGTGATGGAATCCACC	11584
Qy 660	-----	678
Db 11584	TTTCATTAAAGCTGATGATCTCTGTGTGCTCCACTGCTGCTGACCTGGAAG-TGAGT	11590
Qy 678	-----	678
Db 11590	ACAGGTCAGCTCACCCACTATACCCCGTGTCTGACGCCACAGCTGTACACAGGGCT	11596
Qy 678	-----	678
Db 11596	GCACGTGGCAAGGGGACCCCTTTGTGCAAAAGGAGCAATGACAGTTCACACTCCACGTTG	11602
Qy 679	-----	696
Db 11602	TCCAGTGTCAACGACAGATACCTGGGTGGCCCGCAATGTCTCAGAGGCTCAGAC	11608
Qy 696	-----	716
Db 11608	ACTGCGCTTGAGGCTGACAGAGTGGGCTTTGGAGCCACACAGTGGAGAGTGTACAGCG	11614
Qy 716	-----	728
Db 11614	GGGACCGGCCACATCACCCTGTGAGCCCGCAAGTCAGAACCGGTGTCGGGCT	11620
Qy 728	-----	728
Db 11620	TCTGGCCCCGTGAGGGGAGGGGATATCTGAGGGTAGGTTAAATCCGACGTGACA	11626
Qy 729	-----	738
Db 11626	TGCGAACACCGAAGTGAACCTGCTCCCTTAAAGCTGAAATCTTATGAGGACGAGGAC	11632
Qy 738	-----	758
Db 11632	AACTCCCGGCTCAATATGACAGAACTGGCGGACTTATTCACACAGTCAATGCCATATG	11638
Qy 758	-----	778
Db 11638	ATTCGTGGCGGCGCCCTCTCTCCGAGCATCTCAGAGACCTCAACGGCTCATTCAT	11644
Qy 778	-----	778
Db 11644	CAGGTGCCGAGAGGGGGGCGCGCAGGGGCGAGAGGGTCCCTGAGCTGGGTCCAGAA	11650
Qy 778	-----	778
Db 11650	GTTTGCTGTTTACAGATCAGCTGAACCTAGACAGCTGCTGCAACAACTAAGACCC	11656
Qy 778	-----	778
Db 11656	GATCTGATTTGTTTGCAGTGAAGCCAGGTTGTGTTTACAAATGATGCCCAAACTAG	11662



QY 778 ----- 778  
Db 116627 CATCTCAGGCTGATTTCTCCCTGCTTCCACCATCATGGGTGCGGATTGGTTTATTT 116686  
QY 778 ----- 778  
Db 116687 TGTATTCTTAATTTTGTGCACAGATGTTCTCATCTGAGGGATTCAGATCAGAGACT 116746  
QY 778 ----- 778  
Db 116747 GGCCTCAGTGCATGCTGTCTCTGCGGGTCCCTTTGTCTGAGCCCTTGGGTCACA 116806  
QY 778 ----- 778  
Db 116807 CAGAGTGCCTTTGCCGGACAGTGTCTCTCATTTCCCATTAAGTACTGTGGGCG 116866  
QY 778 ----- 778  
Db 116867 TCTCATTAACGACAGAGAGACAGGTTTCAGTGTGATTTACACCAATTCACTTGGG 116926  
QY 779 ----- 779  
Db 116927 ATTGAGGACGAGACTTACCACTTACCCACTTCTCAGCTGATCCCGCTTACATTTACG 116986  
QY 781 ----- 781  
Db 116987 AGCTCTCTCAGACCCACACACCTGGCATCCGAGTCTCTGAGATGAGCTGTGC---GSGT 117043  
QY 801 ----- 801  
Db 117044 GCGCCAGCTCTATGCTGCTGCAGAGACCCCGCATATTTCAGAGAGACACCTTAAGTACA 117103  
QY 821 ----- 826  
Db 117104 TCTCTTGTCTGGGCAAGGTAGTGGCGGCGCATGTGGGGAGGAACGTGGTGGTGAT 117163  
QY 827 ----- 834  
Db 117164 GGGTCAGGTGAGACCTGCTCTCATCTTCATCCACAGGGCAACTTGGCGAGGAGCTG 117223  
QY 835 ----- 854  
Db 117224 TGCCGCTATGACCCCTGGGGGACATACGGAACCTGCTGGCAGTGAACAGCTACAG 117283  
QY 855 ----- 874  
Db 117284 CACAGCGGGCCAGACACAGCAGAGGAGCTTCCAGCGGAGATTGATCTTAAGGCTCTG 117343  
QY 875 ----- 889  
Db 117344 CACAGCGACTTCATGTCATGAGTACCGGGAGATCAGCTATGGCCAGGTGAGCGCAGCAG 117403  
QY 889 ----- 889  
Db 117404 CATCTCGGAGAGGGGTGAGATCCGCTTACCTTTCCAGTAGGGCCCTGTGACACA 117463  
QY 889 ----- 889  
Db 117464 AGGTCTTAATCTCCATGATCCAGCTTCTATAGCTGAGGTGAACGATCATACAC 117523  
QY 890 ----- 896  
Db 117524 CCAGTAGAAGGTCTCGAGAGCTACTGTGCTTCCGAGGTCCGACAGGCTTGGGTTG 117583  
QY 897 ----- 915  
Db 117584 GTATGAGTACCTGCCAGCGGCTCTCTCGAGACTTCTCGAGCGCATTCGCGCGGCG 117643  
QY 916 ----- 930  
Db 117644 CTGCACACGACGCGCTACTGCTGTCTTGGCAGATCTGCAAGGTGCGAGCATCGCG 117703  
QY 930 ----- 930  
Db 117704 AGATCGGATGGGTCTCTGGGGTCCGGATCCCGGCGCAGCTGCTATCATGCTCTCT 117763  
QY 931 ----- 949  
Db 117764 GCAGGCAATGAGTACTGGGTGCGCGCGCTGCTGACACCGCTGACCTGCGCGCACA 117823  
QY 949 ----- 969  
Db 117824 CATCTTGTGAGAGCGAGCTCATGTGAAGATGCGGAGCTTGGCTGCTGCTAGCTGCT 117883  
QY 969 ----- 988  
Db 117884 GCGCTGGAAAGACTACTGCTGCTGCGGAGCCTGCGCAAGCCCATCTTTTGGTA 117943  
QY 988 ----- 988  
Db 117944 CAAGACCAAGACTCCAGAGTACGAGCCCACTGAGCCCACTCCGCACTACAGCTT 118003  
QY 988 ----- 988  
Db 118004 AAACGGTCCAGACCCAGCTCCACCTAGTACTTGGCTCGCAAGGCTTGGCCACA 118063  
QY 988 ----- 988  
Db 118064 GCAGGTCAAGCCCTATAGCCAGATCCAGAACCCATGCTCACAACCTACAGTATAC 118123  
QY 988 ----- 988  
Db 118124 CCTCTCACTTACTGACGACTACCCACCCACCCACAGAGTATGCTCAGCTGGTTTA 118183  
QY 988 ----- 988  
Db 118184 GCGACGCCCCCATGTTGCTGCTCATCTCTTGGACAGTCTGCAAGCCCGTACAG 118243  
QY 988 ----- 988  
Db 118244 GTCTCTTTTACTAGCTCGCTCCCTCGGAGGCGCTGCTTTCACGCTATAGGGTCCG 118303  
QY 988 ----- 988  
Db 118304 CCTCTTAGGCTCTTCTACCTTCTGCTGCTCATCTTACTGCTTCCCGTGGCTGCT 118363  
QY 988 ----- 988  
Db 118364 CCCCTTATGACATTATGAGCTATGCTGCCAGGACCTTTCACGCTCTGTCAATCAG 118423  
QY 988 ----- 988  
Db 118424 CTGACCCATGAAGCCCGCTCTTAACACTTATAGTCCCTCTCTCAAGCTGCGCAGCC 118483  
QY 988 ----- 988  
Db 118484 CACAGGTACAGACAGCTGTGAGTGGCTTGTGGGGCGCTGCTATGAGAACTCTC 118543  
QY 988 ----- 988  
Db 118544 TGTTCAGCCCTTCTCTACCCAGAGTGGCCATCTTACATTTAGAGACTGCCCTCCC 118603  
QY 988 ----- 988  
Db 118604 AGGCACTGTTTACACCTGCTCCACTCTCTCATCTGCGCACACACACACTCTGTCCGC 118663  
QY 989 ----- 1006  
Db 118664 TCTCAGTATGCGCCGAGACCTATCTGACACATCTTCTCCGCAATCTGAGCTGTG 118723  
QY 1006 ----- 1026  
Db 118724 GAGCTTGGAGTGTGTGTGAGAGCTCTTACACTCTGCGACAGAGCTGAGCCCATTC 118783  
QY 1026 ----- 1027  
Db 118784 TCTGACACGACGCGCTACTGCTGTCTTGGCAGATCTGCAAGGTGCGAGCATCGCG 117703



Db 118784 CGC-TGTGGGTGCGTCCCATCCCTGCTGCTCTCTGATCTCCAAATCCCTCCT 118842  
 QY 1027 ----- 1027  
 Db 118843 GACCTTACGCCCTATCTGACCCAGCCCTTCTCTGACCTCCAGATCATCTCTGAC 118902  
 QY 1027 ----- 1027  
 Db 118903 CTCAGTCTCCCTCGGGAATCCAGCCCTCTCTCTCATCTCCAGATCCCTTGTGACC 118962  
 QY 1027 ----- 1027  
 Db 118963 CAGCCCTTCCCTGACCCCAAGCCCTTCTCTCATGACGCCCCCTTCTGACCCCTAAC 119022  
 QY 1028 ----- 1027  
 Db 119022 ACTCTGCCCCCTACTGCTCTCTCTCCACAGAGATTCTCTGAGCATGATGGGCGCTGAGCG 119082  
 QY 1037 gclugllyProProLeuCySarqLeuLeuGlulLeuLeuLagluLglArqArqLeuProPr 1057  
 Db 119083 TGAAGGACCCCGCT 119142  
 QY 1057 gProProthrcySPProthrglu----- 1064  
 Db 119143 ACCTCCCACTGCTCCCGCCAGGAGTGAGAGAGAGAGACTCAAGTTTCTCAGTTTCAATTC 119202  
 QY 1064 ----- 1064  
 Db 119203 TGATTGACAGAGCGGCTCTGACTGGGAATACCTTACAGATGAGGAGCGCTACGCTTC 119262  
 QY 1064 ----- 1064  
 Db 119263 AAGGCTTCAGAAATGTGTTTTCACCTTTGTTGAGCAGGCTCTGATGTAGCCTTG 119322  
 QY 1064 ----- 1064  
 Db 119323 TCTGGCTGGAACTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 119382  
 QY 1064 ----- 1064  
 Db 119383 ATGTGCACACACACACCCCACTCAATTTTCAATTTTCTTTTATTATTACTTA 119442  
 QY 1064 ----- 1064  
 Db 119443 TTTATTTTACAGTCTGAGATTTTGAAGCAGGCTTACTCATGCTAGGCAAGCAGCG 119502  
 QY 1064 ----- 1064  
 Db 119503 ACCCTGTGTCATGATTAATTCGGGGGATGGGGGTTGGGGAAGGAGCGCGCTTCCA 119562  
 QY 1064 ----- 1064  
 Db 119563 GGAATGATGCGCCCTATTGAGAGCCAGAGAGGTCAATTCCTAGTAGACATGCT 119622  
 QY 1064 ----- 1064  
 Db 119623 GTGGGCTGGGGTGAAGAGAGAGATGTCCGGCAGATGGGGATGTGCTGTC 119682  
 QY 1064 ----- 1064  
 Db 119683 ACCCTAGCATCTGAGCGCCAGGAGAGTGTGCCACAGCAAACTTATTGCTGTC 119742  
 QY 1064 ----- 1064  
 Db 119743 AGAGATCCATTAAGAGACCATGAGAGCTGTGAGAAAGTAAAGGGCCATTAACTGAGT 119802  
 QY 1065 ----- 1067  
 Db 119803 GGAAMACAGCCTCTCAGCTCAGGACCCCACTTGTCTTCCACACAGAGTTCCAGGA 119862  
 QY 1067 uleuMetglLeuCySTrPalAProgluProhIsaspargProAlaphaIaThLeuSe 1087  
 |||||||

Db 119863 GCTCATGACGTGTGCTGGGCGCCAGCCGACAGACGGCCAGCCCTTGGCAACCTGAG 119922  
 QY 1087 rProgluLeuaspProLeuTrpArglyArgProgly 1099  
 Db 119923 CCCCAGCTGAGCGCGCTGTGGCTGGAGAGACCCGGA 119959  
 RESULT 9  
 AC073750  
 LOCUS  
 DEFINITION  
 AC073750 223734 bp DNA linear HTG 29-JUN-2000  
 Mus musculus clone RP23-330D8, WORKING DRAFT SEQUENCE, 28 unordered  
 pieces  
 AC073750  
 AC073750.1 GI:8810367  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
 Mus musculus.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 223734)  
 DOE Joint Genome Institute.  
 Sequencing of Mouse  
 Unpublished  
 2 (bases 1 to 223734)  
 DOE Joint Genome Institute.  
 Direct Submission  
 Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 -----Genome Center  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: <http://www.jgi.doe.gov>  
 -----  
 Project Information  
 Center Project Name: 1863672  
 Center clone name: RPCI-23\_330D8  
 -----  
 Summary Statistics  
 Consensus quality: 196566 bases at least Q40  
 Consensus quality: 208241 bases at least Q30  
 Consensus quality: 210965 bases at least Q20  
 Estimated insert size: 208000; agarose-fp estimation  
 Estimated insert size: 221034; sum-of-ctigs estimation  
 Quality coverage: 8.01 in Q20 bases; agarose-fp estimation  
 Quality coverage: 7.54 in Q20 bases; sum-of-ctigs estimation  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 28 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1 1027: contig of 1027 bp in length  
 \* 1028 1127: gap of unknown length  
 \* 1128 2161: contig of 1034 bp in length  
 \* 2162 2261: gap of unknown length  
 \* 2262 3709: contig of 1448 bp in length  
 \* 3710 3809: gap of unknown length  
 \* 3810 4864: contig of 1055 bp in length  
 \* 4865 4964: gap of unknown length  
 \* 4965 5985: contig of 1021 bp in length  
 \* 5986 6085: gap of unknown length  
 \* 6086 7691: contig of 1606 bp in length  
 \* 7692 7791: gap of unknown length  
 \* 7792 10317: contig of 2526 bp in length  
 \* 10318 10417: gap of unknown length  
 \* 10418 11756: contig of 1339 bp in length  
 \* 11757 11856: gap of unknown length  
 \* 11857 13223: contig of 1367 bp in length  
 \* 13224 13323: gap of unknown length  
 \* 13324 16214: contig of 2791 bp in length  
 \* 16215 16215: gap of unknown length  
 \* 20778: contig of 4564 bp in length



Db 125538	CTGGAGCGGCTACATCCAGGGGCGACACACCGAGACCTTCGTTGGGCTCCCGGGCC	125597
Qy 262	GLNGIUGIuProGlyLeuLeuArGyAlaGlyAspAsnGlyIleProTrpSerAsn	281
Db 125598	CAGAGAGAGCGCGGCTTCGTGGGTGGCGGGGACACAGCATCTCCGTGAGCTCCGGG	125657
Qy 282	AspGlu-----	283
Db 125658	GACCA-CGAGGTCTGGGGCTGGGCTTGAGAGCTGGGTCCGGGGAGCGTGGCGGT	125716
Qy 283	-----	283
Db 125717	TAGTAAGAAGCCCGTCCGACACATCAGGAAAGTGGCTAGTCGGGGGGAGCTGATTGCA	125776
Qy 283	-----	283
Db 125777	CTACGGGGCAGGGGGGGGGGCTCAATTAAAGTGGGTCCGGGATGAGCTACCCAGGA	125836
Qy 283	-----	283
Db 125837	CTGACCTGGGGGCGTGACTAGAAAGGGTGAAGACTAGATAGGGGGGGGACGACAGAA	125896
Qy 283	-----	283
Db 125897	GGGATAGATGATGGATTGGATCGGTGGGTGTGACCGGACGAGCGCTGTAGTCCCTAAC	125956
Qy 283	-----	283
Db 125957	CTTCGTCGCCCGACCCCATCCACCCCCCTCCCCCGTAAGTGAGGTGGGAAAGAGAAG	126016
Qy 283	-----	283
Db 126017	AAGGAGAGCTGTGTGGGTGAGGCTGTATCTCTGGCTCTGCTCTCTTTCT	126076
Qy 283	-----	283
Db 126077	CTCTTTTTTTTAAATGTTTTGCGAGACAGGTTTCTCTGTATGCTCTGGCTGTCTGGA	126136
Qy 283	-----	283
Db 126137	ACTCACTTGTAGACACAGGCTGGGCTGGAACTCAGAAATCCGCTCTGTGCTTCGA	126196
Qy 283	-----	283
Db 126197	GTGCTGATTAAGGGGTGGGCGCACATACACCAGCTACTGCTGTCTTCTTCCTCAG	126256
Qy 284	LeuPheGluThrPheCysAspPheProGluIleValAspValSerIleAsnGluAlaPro	303
Db 126257	CTTTTCCGACCTCTGTGATCTTCCGGAATGTGTGATGTACGATCAAGAGGCGCCCA	126316
Qy 304	ArGyAlaGlyProAlaGlyGluIleAspArGleuValThrValThrArgMetAspGlyHisIle	323
Db 126317	CGTGGGTGGCGGACGAGGAACACCGGCTGTCAACGATGGACGGCCACATC	126376
Qy 324	Leu-----	324
Db 126377	CTGGGTGCTTGGCACCCGTTATCTGTGTCTGATCTGGGAAAGGGGTGGTCCGTT	126436
Qy 325	-----GluAlaGluPheProGlyLeuProGluAlaLeuSe	336
Db 126437	GTTCCTGAAGATCTCTCTCCGACGAGACCGAGGTTTCCGGGGCTGTCCGAGCGCTGTTC	126496
Qy 336	rPheValAlaLeuValAspGlyIlyrPheArGleuIleCysAspSerArgHisIlyrPheCys	356
Db 126497	CTTGTGTGGCCCTGTGTGATGGTACTTCCGCGCATCTCGACTCCAGCATTAATTATCTG	126556
Qy 356	sLySGluValAlaProProArGleuGluGluGluAlaAspValCysHisGlyProIle	376
Db 126557	CAGAGAGGTGGCGCGCGCGGCTCTGTGAGAGAGGAGGCGAGCTGTGCATGTGACCAT	126616
Qy 376	erh-----	377

Db 126617	CACCTTGGAGCTAGTTGGGGTCCAGTGGGGATGGGGCTGGCAGCGAGGACTGTGG	126676
Oy 378	-----LeuAspPheAlaIleHisLysLeuLysAlaA	388
Db 126677	CTTGGACCTGGCTGCTGCTCTTGGCCTAGGTAGACTTTGGCCATCCACAAAGCTGAAGCCG	126736
Oy 388	IaGISeLSeLSeuProGIyThrTyLleLeuArGArSerProGIaAsPtyrAspSerPheL	408
Db 126737	CTGGCTCCCTCCAGGACACTCATATTCTCCGCCGAGCCGCGAGAGCATATGACAGCTTTC	126796
Oy 408	euleuPrrAlaCysVal-----	413
Db 126797	TTCTTACCGCTGGCGT-CCAGGTTGGTTTACTGTGTCGTCAGAGATGGCTTCTGGAGAA	126855
Oy 413	-----	413
Db 126856	AGTGACACCTGAGCTGGAGACAGAAGAGGCGTTTCTATTTTTCTTGAAGACACTGGCC	126915
Oy 413	-----	413
Db 126916	TGTAACTCTAGATTAAAGAGACTGAGGACGAGATCAAGTTTACAGGCTGCTGAGCTAT	126975
Oy 413	-----	413
Db 126976	ATAGAGAAATTAATGTTTCTCATGAGAGGAAAGTTTTTTTCCACAAAGCTGAGGGTCT	127035
Oy 413	-----	413
Db 127036	CTTTCATCCATGGCATTCACATGCATGAGAGGACAGAACTGACACCCATAATCGCTGTG	127095
Oy 413	-----	413
Db 127096	ACCTCCAGAGTGCACGGTGACTTGCATCTACACAGAAACACACATAGATGTAATTT	127155
Oy 413	-----	413
Db 127156	TAAGAACCTGAAAGAGCTCAGGAAACAGTCTCAGCATGCGGGGGGGGGGGAGAG	127215
Oy 413	-----	413
Db 127216	CATCCTTAAAGGTTGGAGACAGAACTGAGCATACGTTCTCCAGAAAGACGAAGGGGT	127275
Oy 414	-----GlnThrProLeuGlyP	419
Db 127276	GGGTAGCTGTGCTCCACATCTGGCTGAGTGCAGAGTCTTGTTCGAGACTCCCTTTGGGC	127335
Oy 419	roAspTyLysGIcYsLeuIleArgGlnAspProSerGIaAAsPheSerLeuValGIYL	439
Db 127336	CGGACTTAAAGGGCTGCTCTATCCGCCAGGAACTCCAGCGGGGGTTCCTCCGTGGTGGCC	127395
Oy 439	eUSeRGIaPProHisArgSerLeuArgIuLeuLeuAlaAlaCysTrpAsnSerGIYLLeuA	459
Db 127396	TCAGCCACACCCACAGAAAGCTCGCGGAGCGCTTCACAGCTGCGGAATCTGGGGTGC	127455
Oy 459	IgValAspGIaAlaAlaLeuTyLLeuThrSerCysAlaProArgProLys-----	476
Db 127456	GAGTGAAGCGGTGCTGCCCTGACCTAACAATCTGTGCTGCCCTCCAGACCCAAAGGTGAGT	127515
Oy 476	-----	476
Db 127516	CAGTTCTCCCTTAACCAAGCCGGCAGACTTTGGGTGGTAGCAGAAACCAAGACCAATGG	127575
Oy 476	-----	476
Db 127576	CGATTAGACAGAGAACACGACTTTTTTGGCACAGTCTTCAGACATCAAGTCAGTTATTTT	127635
Oy 476	-----	476
Db 127636	CAGTTGGCCGCCGTTTCTTCTCCATGGGTGAATGCCACTTAAGCGCAACCCCTGCTGGAGCA	127695
Oy 476	-----	476
Db 127696	CATCTCCCTCTGTGGTGTCTGGCTTAACAAAGACACAGGGGGCTTGAGAGGAACCCAGT	127755

OY	476	----	476
Db	127756	ACGTTCAATGTTCAATCAGCCCCAAGTCAGATCTCAGAGGACTTCCCGGGGGCC	127815
OY	476	----	476
Db	127816	CCAGGCTAAGTATGAGGAGTCGAATCTTGGAACCTTGGCTCTGACTGCTTTCTT	127875
OY	477	-----GluYSerAsnLeuIleValArgArgIlyAsnProIleProIle	494
Db	127876	GCCATCCAGAAAGTCCAAATTTATGTTGGTGGAAAGGGGCTGCACCCCCGCTGGCC	127935
OY	494	roglYysSerProSerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProIle	514
Db	127936	CTGGCTGCTCCCGTCTCTCTGTTGGCGGTGACACAGCTGAGTCTCCACAAATTCACAGG	127995
OY	514	spSerLeuGluTrp-----	518
Db	127996	ACACGCTGAGTGGGTGAGGGTCCCCAGGCTGAGAGCCAGGGCGGGCCCTTTGGAGAG	128055
OY	519	-----HisGluAsnLeu	522
Db	128056	ACGTGGAAGCCACTTCACTGTCCTCGCTGGGTCTCTCCCGCAGACAGAGAACTTG	128115
OY	523	GlyHisGlySerPheThrLysIlePheArgIlySerArgArgIlyValValAspGlyGlu	542
Db	128116	GGTCACGGTCTTTTACCAGATCTTCCTCGCGCAGAGGGGGAGGTCTGTGATGTGAG	128175
OY	543	ThrHisAspSerGluValIleLeuLysValMetAspSerArgHisArgAsnCysMetGlu	562
Db	128176	ACACATGACTCGGAAGTCCTCTGAAGGTCTATGACTCCAGACATGGAACCTGCATGGAG	128235
OY	562	-----	562
Db	128236	GTGCGACGGGAGCAGAGAAAGCCCTTGACAGAACTCATGSACTGGGACCTGGCAGAGG	128295
OY	562	-----	562
Db	128296	GTTACTGGAACCTTAGAAGTCAACACACAGACTGGGGGACATAGTGGGACCTCAAAAC	128355
OY	562	-----	562
Db	128356	AAAGTAGAGTTAGCCATGCTGAGCTGGTTGTAAACACAGACATGAGAGAGCTGGGAGG	128415
OY	562	-----	562
Db	128416	AACTGCT	128475
OY	562	-----	562
Db	128476	CTCTCTCTCTGT	128535
OY	562	-----	562
Db	128536	TGATGTGAGCATGACATCTGTAAAGCCAAAGGACAGGACCAAGGAGGTGTAGAGTTCA	128595
OY	562	-----	562
Db	128596	CTCTCCCCCAAGCCCATCACAGCATGAGAGACTTGAAGTACCTTCTAGTGGACTTCTGA	128655
OY	562	-----	562
Db	128656	CCCACTCTGGGCTAAGTGTGGAAACCAAGTGGGAGAGAAATCATATGTGCTCTCTCTCC	128715
OY	563	-----SerPheLeuGluAlaIleSerLeuMetSerGlnValSerIlyProHisLeuVal	580
Db	128716	CCTTAAGTCTTTCTGTGAACCCCAAGCTTGATGAGCCAAATTCATCCCGCACTGTGTG	128775
OY	581	LeuLeuHisGlyValCysMetLacIlyAspSer-----	591
Db	128776	TTATGTGACGGGCTCTGTGCTATGGCTGTGAGACAG-TGAGCCCGGGGACACACCACGCGCAG	128835

QY	591	-----	-----	-----	591
Db	128835	CCCTGGCCTCTCTCCACACCTGCTTACCTGATGCCTCTGATACCTTAAAGATCATCAC			128894
QY	591	-----	-----	-----	591
Db	128895	CCCGTAGCACCACTACATCCAGAGGCCACAGCCGGGCAACAGATCATTTCCACC			128954
QY	591	-----	-----	-----	591
Db	128955	TTGCACGTAATTGCTATGCAACATTCCCTCCAGCCCAACCCCGTGCCACCCCTCTGTA			129014
QY	591	-----	-----	-----	591
Db	129015	ATCTTGCGCTGATATCTATTCACTTAGCTGATAGGAAACAGAGTCCAGCCTGTCCAAAGA			129074
QY	591	-----	-----	-----	591
Db	129075	GCAGAGGAGAAATGGGGGAGCGCTGCTGAGTACCTGTCACCTGAATACCAACCAACCC			129134
QY	592	-----	-----	-----	592
Db	129135	ACTCTCCCCCTGTCCCGCAGCATATGCTGCGAGGAATTTGTATCTTAGAGCAATGTAC			129194
QY	605	MetTYrLeuAArgLYsArgGLYHisLeuValSerAlaSerTrpIysLeuGlnValThrIys			624
Db	129195	ATGTACCTGCGCAACCTGTGGCCACCTGGTGTCCACCCAGCTGGAACCTGCAGGTGACCAAG			129254
QY	625	GlnLeuAlaTyrAlaLeuAsnTyr-----			632
Db	129255	CAAGCTGGCATATGCCCTTAAGTACTACTTGGTGTGAGTGTCCCTGTGTCATTAAGAGAGTTCCT			129314
QY	632	-----	-----	-----	632
Db	129315	CTATGATGGGGAATGAGAGTAGAGTTCTGTATTGGATAGAGCTTGCTTAAGGACACA			129374
QY	632	-----	-----	-----	632
Db	129375	AGCAACGTTCCAGATGACGTGAGCAGGCTCAGAGGGGCTGTAGTGTCAAGGAGAGCA			129434
QY	632	-----	-----	-----	632
Db	129435	CTGTGAACCATAGGTGTCTTAGGAAGCAGTAGAGTAGAGAGGCCAGGCTAGGGGGTGGG			129494
QY	632	-----	-----	-----	632
Db	129495	CGGGGTGTAGAGATGGAACCTGCTGAGAAGAGGGGGTGTCTTGAATGGCGGATGTGGGG			129554
QY	632	-----	-----	-----	632
Db	129555	CTTTAACTGTCCTTAAGCTGAGTGGGAAGGTAGAGAAAGGGGTGTGCTTAAAGATGGG			129614
QY	632	-----	-----	-----	632
Db	129615	TATGTCAGTGTAAAGGGAGGTTTAGGGTTACACAGCATGACTTCAACGGGCTTACAGTGGT			129674
QY	632	-----	-----	-----	632
Db	129675	TGCGAAAGCCTGTTGTGTGTAGGACAGCATCAGATAGCCTTTGAAGGATGCAGAGAGTTT			129734
QY	633	-----	-----	-----	640
Db	129735	GCTCACAACTACTCCCTGTCCTCTCTGTAACACTTAGAGAGGACAAAGGCTTCTCTCA			129794
QY	640	sgLYAsnValSerAlaArgLYsValLeuLeuAlaArgGLYgLYAspGLYAsnProPr			660
Db	129795	CGGGAAGCTTACGACGAGAGGTGCTCTGCTGTAGAGGGGTGTATGGGAATTCACC			129854
QY	660	oPheIleLYsLeuSerAspProGLYValSerProThrValLeuSerLeuGlnMet-----			678
Db	129855	TTTCATTAAAGCTAGTGATCTGTGTGTGTCAGTCCCACTGTGTCAGACCTCGAAAG-TGAGT			129913
QY	678	-----	-----	-----	678

Db 129914	ACAGGCTAGCTTACCCCACTATACCCCGTGTCTGACAGCCACAGTGTCTACACAGAGGCGCT	129973
Oy 678	-----	678
Db 129974	GCACGTGGCAAGGGGACCCCTTTGTGCMAAGGTGACCATGACAGTTCTCATTCCACAGCTTC	130033
Oy 679	-----LeuThrAspArgLleProTrpValAlaProGluCysLeuGlnGluAlaGlnThr	696
Db 130034	TCGAGTCTCTACCCGACAGAAATACCTCGGGGTGGCCCCCAATGTGCTCCAGAGAGCTCTGAC	130093
Oy 696	rLeuCysLeuGlnValAlaAspLysTrpGlyPheGlyValThrThrTrpGluValPheGlnAr	716
Db 130094	ACTGCGCTTGAGAGCGTACAGAGGGGCGCTTTTGAGGCCACACGTTGGAGGTTCACGCG	130153
Oy 716	gGlyProAlaHisLleThrSerLeuGluProAla-Lys	728
Db 130154	GGGACCGCGCCACATCATCCTCGCTGAGCGCGCCAAAGTCAGAACCGGTGTCTCGGCGCT	130213
Oy 728	-----	728
Db 130214	TCGTGGCCCGGTGAGGGGAGGGGGATATCTGAGGCTAAGGTTAAATCGGACCTGAGCA	130273
Oy 729	-----LysLeuLysPheTyrGluAspGlnGlyG	738
Db 130274	TCGCCAACCCGAAGTGAACCTCGTCCCTTAGAAGCTGAAGTTCTATGAGAGACAGAGGAC	130333
Oy 738	LeuLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuLleThrGlnCysMetAlaTyrA	758
Db 130334	AGCTGCGCGCTCTCAAAATGACAGAACTGGGGGACTTATCACAGGTCAATGCGGCTTAG	130393
Oy 758	sProGlyLysArgTrpProSerPheArgAlaLleLeuArGAspLeuAsnGlyLeuLleThrs	778
Db 130394	ATCTGTGGCGCGGCCCTCTCTTCGAGTATCTCAGAGACGTCACAGGCGCTCATTACAT	130453
Oy 778	er-----	778
Db 130454	CAGGTGGCGGAGAGGGGGGGCGGCGAGGGCGGAGAGGCTCCTGAGTGGGTCCAGGAA	130513
Oy 778	-----	778
Db 130514	GTTTCTGTGTTCTAGCAGTCAAGCTGAACCTAAGACAGCTGTGCAACCAACTAAGACCC	130573
Oy 778	-----	778
Db 130574	GATCTGCAATTTGTTTGTGACGTGTAAGCCAGGTTGTGTTTACAAATGTAAGCCAACTAG	130633
Oy 778	-----	778
Db 130634	CATCTCAGGCTGATTTTCTCCGCTTCAACATCATGSGTGTGGGATTTGGTTTATTT	130693
Oy 778	-----	778
Db 130694	TGTTATTTCTTAATTTTGTGACAGAAGATGTTCTCATCTGAGGGTATCCAAATCAGACACT	130753
Oy 778	-----	778
Db 130754	GCGCTCACTGCATGGCTGTGTCTCTCGGGTCCCGCTTTTGTCTGTAGCCGTTGGSGTCCA	130813
Oy 778	-----	778
Db 130814	CAGAGTGCCTTTGGCCCGGACAGTGTGCTGTCTATCCCAATTAAGTACTGTGGCG	130873
Oy 778	-----	778
Db 130874	TCCTCATAGCGACAGAGAGAGCAAGTTTCAAGTGAAGTTTACACACCAATTCAGTTTGGG	130933
Oy 779	-----AspTyrG	781
Db 130934	ATTGAGCGCAGAGACTTACACTTACCCCACTTCTCACAGTATCCCGCGCCTTGAGTTAGC	130993
Oy 781	LeuLeuSerAspProThrProGlyLleProSerProArgAspGluLeuCysValAlaG	801

Db 13094	ACCTCTCTCAGACCCCAACACTGGGCTCCGAGTCTCTGAGATGAGGCTGTC---	-GGTG	13105
QY	801	LYALAGLLeuTYrAlAcYSGLAsProLaIAlpheLuoLuarGhtsLeuLYsTYrI	821
Db 131051	GGCGCACACTCTATGCTCGCCAGAGACCCCGCATATTTCAGAGAGACACTTAAGTACA	131110	
QY	821	LeSerLeuLeuGLys-----	826
Db 131111	TCTCTTTCGTGGGCAAGTGAAGTGGGGGCGATGTGGGGAGAGAACGTGGGTGGGTGAT	131170	
QY	827	-----GlyAsPheGLyservaGLuLeu	834
Db 131171	GGGTGAGTGGACACTGCTCTCATCTCCACAGGGGCAACTTGGAGGTGGAGCTTG	131230	
QY	835	CysArTYrAsProLeuGLysAsPaNthrgLYProLeuValAlaValLYsGLnLeuGLn	854
Db 131211	TGGCGCTATACCCCTCTGGGGGACATATAGGGAGCCCTGTGGGCACTGMAACACTACAG	131290	
QY	855	HisSerValProAspGLnIAsPheGLnAsPheGLnArgGLuLeuGLnLeuLYsAlaLeu	874
Db 131291	CACAGCGGGCAACACCAACGAGAGGGACTTCCAGCGGGAGATTCAAGATCTTAAAGCTCTG	131350	
QY	875	HisSerAsPheLeuValLYsTYrArGLysTYrArGLysLeuTYrGLYPro-----	889
Db 131351	CACAGGCACCTTCATCGCTCAAGTACCGGGGAGTACAGTATGGGCCAGGTGAGCGGCACAG	131410	
QY	889	-----	889
Db 131411	CATCTGGGAGACGGGTTCGAGTCCGCCCTTACCCCTTCCACGTAGGCCCTGTGMAACA	131470	
QY	889	-----	889
Db 131471	AGGTGCTTAACCTCCATGATCCAGCTTCTATAGCTGGAAGTTGAACCAATCATACAC	131530	
QY	890	-----GlyArgLInSerLeuArgLeu	896
Db 131531	CCAGTGAAGAGTGCCTCGAGAGACTACTGTGCTGTTCCAGAGTGCACAGAGCCTCGGGTTG	131590	
QY	897	ValMetGLuTYrLeuProSerGLYcysLeuArgAsPheLeuGLnArgHisArgLY---	915
Db 131591	GTGATGTGAATACCTCCCGAGGGCTGCTCGAGACTTCCTGCAAGCGCATGGCCGCGC	131650	
QY	916	LeuHisThrAspArgLeuLeuLeuPheAlaTYrGLnLeuLYsLYs-----	930
Db 131651	CTGCACACCGAGCCGCTACTGCTGTGCTGTGGCAATCTGCAGAGGTGGCAGATCCGCG	131710	
QY	930	-----	930
Db 131711	AGATCGGAGTGGGCTCTCTGGGGTCCGGGATCCCGGGCCCACTGCTATAGGCTCTCT	131770	
QY	931	-----GlyMetGLuTYrLeuGLYAlaArgArgCysValHisArgAsPheLeuAlaArgAs	949
Db 131771	GCAGGGCATGGAGTACCTGGGGTGGCGGCTGCCGTGCATACCTGTGACCTGGCTGGCGGCA	131830	
QY	949	nLeuValGLuSerGLuAlaHisValLYsIleAlaAsPheGLYLeuAlaLYsLeuLe	969
Db 131831	CATCTTGGTGGAGAGCGAGGCTCATGTGAAGATCGGGCACTTGGCCCTCGCTAAAGCTGT	131890	
QY	969	uProLeuGLYAsPryTYrValValaArgLInProGLYInSerProIlePheTYr--	988
Db 131891	GCCCTTGGGAAGAGACTACTACGTGTGTCCGCGAGCCTGGCCAAAGCCCATCTTTTGTA	131950	
QY	988	-----	988
Db 131951	CAAGACCAAGACTTCAGAGCTCAGCGCCACTAGGCCCAACTCCCGCACTTCACAGCTTT	132010	
QY	988	-----	988
Db 132011	AAACCGGTCCAGACCCAGGCTCACTCACTAGTACCTTTGCCCTTCGCAAAAGCCTTGGCCACA	132070	
QY	988	-----	988
Db 132071	GCAGGGTCCACCCCTATAGSCAGAGTTCACAGAAAGCATCTGTACAAACTTCACAGATGTACC	132130	

QY 988 ----- 988  
 Db 132131 CCTCTCACTAGTACAGACTACCCACCCACAGAGTGTGCTCAGCTTTA 132190  
 QY 988 ----- 988  
 Db 132191 GCCACGCCCCCATTTGTTGGTCTCATCTCTTTCAGAGTCTGCAAGCCGTCACAG 132250  
 QY 988 ----- 988  
 Db 132251 GTCCTTTTCTAGCTCCGCTCCCGACAGGGCCCTGCTTTCAGCGCTCTATGGGTCGCG 132310  
 QY 988 ----- 988  
 Db 132311 CCTCTTACGCTCTTCTACTCTCCGTCATCTTACTGTCTTCCGCTGCTGCTCT 132370  
 QY 988 ----- 988  
 Db 132371 CCTCTTACATTTACGCTATCTGCCAGGACATTTTCAGCTGCTGTATTCATG 132430  
 QY 988 ----- 988  
 Db 132431 CTGACCCACTCAAGCCCGCTCTTAACACTTAGAGTCCCTCTCTCAGCTGGCCAGGCC 132490  
 QY 988 ----- 988  
 Db 132491 CACAGGTACACAGAGCCTGTAGTCCCTTCTGGGGCCGTGCTAGTAGGAAGTCTCC 132550  
 QY 988 ----- 988  
 Db 132551 TGTTCAGCCCTTCTCTCACCACAGCTGCGCATCTTCAATGAGAGCTCCCTCCC 132610  
 QY 988 ----- 988  
 Db 132611 AGGCCACTGTTTCACACCTGCCCCACTCTCTACTGCGCACACACCTCTGTCGGC 132670  
 QY 989 ----- 1006  
 Db 132671 TCTCAGGTATGCCCCGAGCTCTATGACACATCTTCTCCGCGCATCTGAGGTGTG 132730  
 QY 1006 pserPheglyValIleuLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProse 1026  
 Db 132731 GACCTTCGAGGTGTGTGTGACAGACTTCACTACTCGACAGAGCTGAGCCATTC 132790  
 QY 1026 rAla----- 1027  
 Db 132791 CGC-TGTGCGTGCCTGCCCCATCCCTGGTCTCCCTCTTGTATCTCAATCCCTCCT 132849  
 QY 1027 ----- 1027  
 Db 132850 GACCTTAGCCCTATCTATACCCACGCCCTTCTCTCTAGCTCAGATCATCTCCTGAC 132909  
 QY 1027 ----- 1027  
 Db 132910 CTCAGTCTCCCTCCGGAATCCAGCCCTCTCTCATCTCAGATCCCTTTTGAGCC 132959  
 QY 1027 ----- 1027  
 Db 132970 CAGCCCTTCCCTACCCCCAGCCCTTCTCATGAGCTGACGCCCTTCTGACCTTAAC 133029  
 QY 1028 ----- 1037  
 Db 133030 AGCTCGCCCTAACCTGCTCTCTCCACAGAGAGTCTCTGAGCATGTGATGGGCTGAGCG 133089  
 QY 1037 gGluGlyProPheLeuSerArgLeuLeuGluLeuLeuAlaGluGlyArgLeuProP 1057  
 Db 133090 TGAAGAGACCCCGCTCTGCTGCTCTCTGAGCTGCTGACAGAGGCCGAGCCCTCCACCC 133149  
 QY 1057 oProProThrCysProThrGlu----- 1064  
 Db 133150 ACCTCCACAGCTGCCCCACAGAGGTGAGAGAGGACTCAAGTTTCTCAGTTTCAATTTCT 133209

QY 1064 ----- 1064  
 Db 133210 TGATTGACAGACGGCGTCTGACTGGGAATACITTTACAGGATGAGGAGCGCTCAGCTTC 133269  
 QY 1064 ----- 1064  
 Db 133270 AAGGGCTTCAGAAATGTGTGTTTCACTTTTGTGTGAGCGAAGGTCTTGCTATGTAGCTTG 133329  
 QY 1064 ----- 1064  
 Db 133330 TCTGGCTGGAACCTGACATCTCTCTGCTCAGTTTCTCTGAGATGAGAACAAATGT 133389  
 QY 1064 ----- 1064  
 Db 133390 ATGTGCCACACAGACACCCCACTCCATTTTCAATTTTCTTTTATTTACTTTA 133449  
 QY 1064 ----- 1064  
 Db 133450 TTTATTTTTCAGTGTGATTTGAGACCAAGGCTTACTCATGTAGGCAACGACGCG 133509  
 QY 1064 ----- 1064  
 Db 133510 ACCCTGTCTATGATTAATTCGGGGGATGGGGGTTGGGAAGGAGCGACCGCTTCA 133569  
 QY 1064 ----- 1064  
 Db 133570 GGAATGATGGCCCTATTGGAGAGCCAGAGAGGTCAATCCCACTAGGACATGTGCT 133629  
 QY 1064 ----- 1064  
 Db 133630 GTGGGCTGGGGTGAGTAAGAGAGAGATGTGCGGCGAGATGGGGATGTGTCTGTC 133689  
 QY 1064 ----- 1064  
 Db 133690 ACCCTAGCATCTGACAGGCCAGGCGAGAGTGTGCCACAGGCAACTTATGTTCTGTC 133749  
 QY 1064 ----- 1064  
 Db 133750 AGAGATCCATAGAGAGCATGAGAGCTGTGAGAGAGTAGGGGCCATTAACCTAGT 133809  
 QY 1065 ----- 1067  
 Db 133810 GGAACACAGCGCTCTCAGCTCAGGACGCCGCCACCTTCTTCCACACGATTCAGGA 133859  
 QY 1067 uLeuMetIleuLeuYsTrpAlaProGluProIleAspArgProAlaPheAlaThrLeuSe 1087  
 Db 133870 GCTCATGACGCTGTGTGTGGGCGCCAGCCGACCGCAGACCGGCGACCTTGGCACCTGAG 133929  
 QY 1087 rProGluLeuAspProLeuThrParGlyArgProGly 1099  
 Db 133930 CCCCAGCTGAGCGCGCTGTGGCTGGAGAGACCCGGA 133966  
 RESULT 10  
 MMU71201 10326 bp DNA linear 1 ROD 15-JAN-1997  
 LOCUS MMU71201  
 DEFINITION Mus musculus protein tyrosine kinase Jak3 gene, complete cds.  
 ACCESSION U71201  
 VERSION U71201.1 GI:1778748  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 10326)  
 AUTHORS Kumar, A., Toscani, A., Rane, S., and Reddy, E.P.  
 TITLE Structural organization and chromosomal mapping of Jak3 locus  
 JOURNAL Oncogene 13 (9), 2009-2014 (1996)  
 MEDLINE 97088640  
 PUBMED 8934548  
 REFERENCE 2 (bases 1 to 10326)  
 AUTHORS Kumar, A. and Reddy, E.P.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-SEP-1996) Fels Institute, Temple University, 3307 N.



Db	1422	GTCCGCTCCGACGCCGACCGCTACGCGTCA	TAT-CTGAGACTTGAGCGGC	1480
Qy	245	euHISProIALAIAthrThrGIunhrPhear	gValGIyleuProGIyAlaInGIuIeP	265
Db	1481	TACATCCAGCGGCCACACGAGACCTTCG	TGTGGGCTCCCGGCGCCAGAGAGC	1540
Qy	265	roGIyLeuIeuargValaIaGIyaspasn	GIyIleProTPSerSerAsp-----	282
Db	1541	CGGGGCTTCGTGTGTGGCGGGGACACAG	CGCATCTCTGAGCTCCGGGACCAGAGG	1600
Qy	282	-----	-----	282
Db	1601	TGCTGGGCTGGGCTTGAGAAAGCGGGG	GTCCGGGAGCGCTGGCGGTAGTAAGAAGC	1660
Qy	282	-----	-----	282
Db	1661	CCGTCCGAACTACAGAAAGTGCTAGTCC	CGGGCGGGCTGATTCACIACGGGGCAGG	1720
Qy	282	-----	-----	282
Db	1721	GGGGCGGGGTCAATTTAAAGTGGGTCCG	GGGATGGAGCTGCCCGAGGACTGACTGGG	1780
Qy	282	-----	-----	282
Db	1781	GCGTGACTAGAAAGGTGAGACTAGATAG	GGGGGCGGGGACGACAGAGCGCATAGTGG	1840
Qy	282	-----	-----	282
Db	1841	ATTGAGTCACTGGGCGATGACCCGACCG	ACGCGCTGTCTTACCTCCCTAACCTGTC	1900
Qy	282	-----	-----	282
Db	1901	CCCCCACAAGTGAGAGGTGGGAAGAGA	GAAGAGAGCGCTGTGCTGGGTCAAGTCTG	1960
Qy	282	-----	-----	282
Db	1961	TAAATCCCTGGCTTGTGTCTCTTCTCT	TTTTTTTTTAAATGATTTTTCGAGACAGG	2020
Qy	282	-----	-----	282
Db	2021	GTTCCTCGTATAGCTGTGCTGTCTTGGA	CTCATCTTTGTAAACACAGCTGGCTCGAA	2080
Qy	282	-----	-----	282
Db	2081	CTCAGAAATCCGCTGCTGCTGCTGCAG	TGCTGATTAAGGCGGTGACCTAC	2140
Qy	283	-----	-----	294
Db	2141	GCGGTACTGCTCTCTCTTCTTCTCTGCT	TCAGCTTTTCCAGACTTCTGTCCGAAA	2200
Qy	294	IeValAspValSerIleasnGlnIaProAr	gValGIyProAlaGIyGluHISArgLeuV	314
Db	2201	TCTGTGATGTCAATCAACAGC-CCACGT	GTGGGTCCGGCAGGAG-CACCGGTGG	2257
Qy	314	AlhrValThrArgMetAspGIyHISIleu	-----	324
Db	2258	TCACGTGCACGAGTGGACGCGACATCTG	TGGGTCTTGCCACCCGTGGTTCGCTG	2317
Qy	325	-----	-----	326
Db	2318	CTGATCTCNGAAGAGGGTGGGCTCCGT	TGTTGCTGAAAGACTTCTCCGCMCAGAGCGC	2377
Qy	327	GIuPheProGIyLeuProGIuIaIeUerP	heValAlaIeUValAspGIyTYrPheArg	346
Db	2378	GAGTTTCGGGGCTGCTGAGCGCTGTCTT	CCGTCCGTGAGATGGGTACTTCGC	2437
Qy	347	LeuIleCysAspSerArgHISTYrPheCys	LySGIuValAlaIaProProArgLeuIeu	366
Db	2438	CTGATCTGCAGCTCCAGCGCATTTCTG	CAAGGAGTGGGGCGCCACCGCTGCTGAG	2497
Qy	367	GIuGIuIaAspValCysHISGIyProIeThr	-----	377
Db	2498	GAGAGCGGAGGTGTGCTGACATGACCA	CTACAGCTTAGAGCTAGTTGGGTCAACAGTGG	2557

OY	377	-----	377
Db	2558	GGATGGCGATGGGACGAGCACTGTCCTTGGGTGGACCTGGCTGCTCTTGGCCTAGG	2617
OY	378	LeuaspPhealaIleHisIysIeuIysAlaIaGlySerLeuProGlyThrTyrIleLeu	397
Db	2618	TTTAGCTTTGGCCATCCACAGCTGAAGGCCGCTGGTCCCTCCAGGCACTATATTTCTC	2677
OY	398	ArgArgSerProGlnAspPyrAspSerPheLeuLeuThrAlaCysVal	413
Db	2678	CCCGGAGCGCCGAGAGACTATACAGCTTCTTCTTACCGCTGGCTCAGGTGGTTTA	2737
OY	413	-----	413
Db	2738	CTGCTGTGCCAGGATCGTNTNCCCTGGAGAGAACTGACACCTGAGCTGGGAGCAAAAGAGGC	2797
OY	413	-----	413
Db	2798	ATTTCATATTTTCCTTGAAGAGACTGCTGTAACTAGAAATTTAAGAGACTGAGCAG	2857
OY	413	-----	413
Db	2858	GATCAAGTTTACAGCGCGGCTGAGCTATATGAGAAATTTATTTCTCATGAGGAAAAAGG	2917
OY	413	-----	413
Db	2918	TTTTTCCACCAAGCCGTGAAGGCTGCTTTCATTCATTCATGCAATTCACAGCANTAGAGACA	2977
OY	413	-----	413
Db	2978	GAACCTGACACCCATTAATTCGCTGTACCTCCACACTGCAGGTGACTTGCAATCTACACA	3037
OY	413	-----	413
Db	3038	GAACACACATTAAGTGAATCTAATTTAAGACTCGAAAGCTCAGAAACAGATTCTC	3097
OY	413	-----	413
Db	3098	AGCATGGGGGGGGGAGAGAGAGACATCTTGACGTTGAGACAGACAGCTGTATACA	3157
OY	413	-----	413
Db	3158	GTCCTCCGAAGAAGCAAGAGGGGTGGTAGCTGCTCCACATCTGGCTCAGGTGAGG	3217
OY	414	-----GlnThrProLeuGlyProAspPyrTyrGlyCysLeuIleArgGlnAspPro	430
Db	3218	TCCCTTTTGGACAGACCTCTTGCCGCCACATCAAGGGCTGCTCCTCCTCCTCCAGGAGCC	3277
OY	431	SerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeu	450
Db	3278	ACCGGGGCTTCTCCTCGTGGCTGAGCCTGAGCAGCCCAAGAGCCGCGGAGCTGCTT	3336
OY	451	AlaAlaCysTyrPasnSerGlyLeuArgValAspGlyAlaIleLeuTyrLeuThrSerCys	470
Db	3337	GCAG--TGCTGGAATTTCTGGGCTGCGAGTAGACGGTGGTCCCTGAACCTTAACCTCTGC	3394
OY	471	CysAlaProArgProLys-----	476
Db	3395	TTCGCTCCAGACCCAGGGGTAGTCAGTNCCTCCCTAACAGGCCCGAGACTCTTGGGT	3454
OY	476	-----	476
Db	3455	GGTAGAAGAACAGAGACAGCCAGTCGATAGACAGAGAACAGACTTATGGGACAGCTTTC	3514
OY	476	-----	476
Db	3515	AGAAATCAAGTCAATTATGTGCAAGTTGGCCGCCGCTTACTTCCATGGGATGCCACTA	3574
OY	476	-----	476
Db	3575	AGCGGAACCCCTCTGGACACATCTCCCTCTGTGTGTCTCTGGCTAACAAAGACCACGG	3634





Db 5794 ACCCACTATACCCCGTGTCTGCAGCCACAGCTCTCAACAGGGCCCTGCACCTGGAAAGG 5853  
 QY 678 -----MetLeuThra 681  
 Db 5854 GACCCCTTGTGCAGAAAGTGCATGACAGTNCCTACCTCCACGCTTCTCCAGTGTCTCACC 5913  
 QY 681 spArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 701  
 Db 5914 ACAGAAATACCTGGGTGGCCCGCAATGTCTCCAGAGGCTCAGACACTCGCTTGAGG 5973  
 QY 701 laAspIlystrpGlyPheGlyAlaThrThrProGluValPheGlnArgGlyProAlaHis 721  
 Db 5974 CTGCAAGTGGGGCTTTGGAGCCACAGTGGAGGThTTCAGCGGGGGACCCGCCACA 6033  
 QY 721 IethrserLeuGluProAlaLys----- 728  
 Db 6034 TCACCTGCTGGAGCCGCCCAAGTCAAGAACCTGTGTGGGTCTCGCCTTCTGGGCCACG 6093  
 QY 728 ----- 728  
 Db 6094 GTGAAGGAGCGGGATATCTGTAGGGTATTAATCTGAGCCGACATCGCAACAC 6153  
 QY 729 -----LysLeuLysPheTyrGluAspGlnGlyLeuPro 741  
 Db 6154 CGAAGTGAACCTCGTCCCTTAGAAGTGAAGTTCTATGAGGACGAGGACAGCTGCCG 6213  
 QY 741 laLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrTrsProGly 761  
 Db 6214 CTCTCAATGTGCAGAACATGGCGGAGCTTATCACACAGTGCATGGCTATCTCTGGCC 6273  
 QY 761 rArgProserPheArgIleAlaIleuArgAspLeuAsnGlyLeuIleThrser----- 778  
 Db 6274 GGGGCCCTCTCTCCGAGTATCTCAGAGACTCAAGGGCTCATTTACATCAGTGCAG 6333  
 QY 778 ----- 778  
 Db 6334 GAGAGGGGGGGGCGAGGGCGGAGAGGTCCTAGACTGGGTCCAGAGATTGCTGG 6393  
 QY 778 ----- 778  
 Db 6394 TTCTAGACAGTCAAGCTGAACATAAGACAGCTGTGCACAAATCTAAGACCGATCTGAT 6453  
 QY 778 ----- 778  
 Db 6454 TTGTTTGCAGTGTAGCCAGGTGGTGTATTACAATGTAGCCAAACTAGATCTCAGG 6513  
 QY 778 ----- 778  
 Db 6514 CTGGATTTCNCCTGCTTACCATCATGAGTGTGGATTGTTTATTATTATTCT 6573  
 QY 778 ----- 778  
 Db 6574 TAAATTTTGCACAGATGTCTCATCTAGATATCGAGTACAGACAGAGAGTCACTG 6633  
 QY 778 ----- 778  
 Db 6634 CATGGCTGTCTCTCGGTTCCCTTTTGTCTAGCCGTGGGATCCACAGAGTGTGC 6693  
 QY 778 ----- 778  
 Db 6694 CTTTCCCGAGACAGTGTCTCATTTCCCATAGTACTGTGGCGTCTCATAGCGA 6753  
 QY 778 ----- 778  
 Db 6754 CAGAGAGCAAGGTTCAGGTGAGTTACACACAAATTCAGTTGGGATTGAGCAGAG 6813  
 QY 779 -----AspTyrGluLeuLeuSer 785  
 Db 6814 ACTTACACTTACCCCACTTCTCAGCTGATCCCGCCCTAGATTAAGAGACTCTCTCAG 6873  
 QY 785 spProThrProGlyIleProserProArgAspGluLeuCysValAlaGlyAlaGlnLeu 805  
 Db 6874 ACCCAACACCTGGCATCCCGAGTCTCGAGATGAGTGTGC---GGTGGCGCCAGCTCT 6930

QY 805 yAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLysTyrIleSerLeuLeu 825  
 Db 6931 ATGCCTGGCAGGACCCCGCATATTGAGAGAGACACACTTAAGTACTCTTTGGCG 6990  
 QY 825 Lys----- 826  
 Db 6991 GCAAGGTGATGGGGGGCATGTGGGGAGAACGTGGGTGGTGGATGAGGTACGTGA 7050  
 QY 827 -----GlyAsnPheGlySerValGluLeuCysArgTyrAsp 839  
 Db 7051 CACTGCCCTTCATCTCCACAGGCAACTTGGCAGCGGTGAGCTGTGCCCTTAGAC 7110  
 QY 839 roLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPro 859  
 Db 7111 CCTG---GACAAATAGGGAGCCCTGTGTGCGAGTGAACAGCTACAGCACAGCGGCCAG 7167  
 QY 859 spGlnIleArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPhe 879  
 Db 7168 ACCAGCAGAGGAGACTTCCAGCGGAGATTCAATCTTAAGCTCTGCACAGCGACTTCA 7227  
 QY 879 leValLysTyrArgGlyValLysTyrGlyPro----- 889  
 Db 7228 TCGTCAAGTACCGGGAGAGTACGCTATGGCCAGGTGAGCGCCAGCANCCTGGGAGCG 7287  
 QY 889 ----- 889  
 Db 7288 GTTCGAGGNNCGCTTACNCTTNNCCAGTAGGCCGTTGAACAGSTCGTTAATCCAT 7347  
 QY 889 ----- 889  
 Db 7348 GAATCCAGCTTCTATAGCTGAGAGTTGAAGCAAAATACCAGTAGAGTGTCTGAGA 7407  
 QY 890 -----GlyArgGlnSerLeuArgLeuValMetGluTyrLeuProser 903  
 Db 7408 GCTACGTGCTGTTCGCGAGTGGCCAGAGCTCGGTGGTGTGATGAGTACCTGCCACG 7467  
 QY 904 GlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHisThrAspArgLeu 922  
 Db 7468 GGTGCTGGAGACTTCTCTGACAGGCATCGCGGCCCTTGACACACCGCCCTACTG 7527  
 QY 923 LeuPheAlaTrpGlnIle----- 928  
 Db 7528 CTGTTCGCTTGGCAGATCTGCAGATGCGACATCGCGAGATCGGGATGGGTCTGGG 7587  
 QY 929 -----CysLysGlyMetGluTyrLeuGly 936  
 Db 7588 GTCGGAGTCCGGGCCACAGCTGATCATGCTGTCTGAG--GGCATGGAGTACTGGGT 7646  
 QY 937 AlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAla 956  
 Db 7647 GCGGCCCGCTCGTACACCGTGAACCTGGCTGCGGCAACATCTTGGTGGAGAGCGAGCT 7706  
 QY 957 HisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyr 976  
 Db 7707 CATGTGAAGATCGGGACTTGTGGCTCTGCTAAGCTGTGCCCCCTGGGAAAGGACTACTAC 7766  
 QY 977 ValValArgGluProGlyGlnSerProIlePheTrpTyrAla----- 990  
 Db 7767 GTGTCCGGAGCCGTGGCAAGCCCATTTTGTGTACAA--GACCAAGACTCCAGAGCT 7825  
 QY 990 ----- 990  
 Db 7826 CAGGCCACTGAGCCCAACTCCGCACTTACAGCTTAACCGGTCCAGACCCAGCTC 7885  
 QY 990 ----- 990  
 Db 7886 CACTTCAGTACCCCTGCTCGCAAGGCTTGCACAGCAGGCTCAGCCCTATAGCAC 7945  
 QY 990 ----- 990  
 Db 7946 GATCCAGAACCAAGCTCACTACAGATGTACCCCTCTCAACTTACTGACAGACTA 8005

Db	9085	ATACCTTTACAGATGATGAGGACGCTCAGCTTCAAGGGCTTCAGAAATGTGTTTCACCTT	9144
QY	1065	-----	1065
Db	9145	TGTNAGAGCAAGGTCTGTGCTATGTAGCCTTGTCTGGCTGGAACTCGACATCTCTGCTCC	9204
QY	1065	-----	1065
Db	9205	TCAGTTTCTCCTGGACTGAGAGAACAAATGTATGTGCCACCAACAACCCCACTCC	9264
QY	1065	-----	1065
Db	9265	ATATTTCAATTTTCTTTATTATTATTACTTATTATTATTTTACAGTGTGAGATTTGAG	9324
QY	1065	-----	1065
Db	9325	CCAGGGTCTTACTCATGCTAGGCAAGCAGCACCCTGTGTCATGATTAATGGGGGATGG	9384
QY	1065	-----	1065
Db	9385	GGGTTGGGAAGGAGGCGCGCCTTCAGAGTGAATGCCCTATTGAGAGCAGCAGG	9444
QY	1065	-----	1065
Db	9445	AGAGTCATTCCAGTAGAGACATGCGTGGCTGGGGTGGAGTAAAGAAAGACAGATG	9504
QY	1065	-----	1065
Db	9505	TGCGCGCAGATGGGGGATGTGTGCTGTCAACCTAGCCATCTGACGCCAGGCGAGAGT	9564
QY	1065	-----	1065
Db	9565	GTTGCCACAGCAACTTACTTCTGTTCAGAGATCCATTAGAGCCATGAGAGCTGTGG	9624
QY	1065	-----	1065
Db	9625	AGAAAGTAGNNNNCCATTAACTGAATGAACAACACTCTCAGCTCAGGATCCGCCA	9684
QY	1066	-----	1066
Db	9685	CCATTGATCTCCACCACACAGGTTCAGAGCTCATGAGCTGTGCTGGCCGAC	9743
QY	1077	OHASAPATGPTCAlaPhaLeuLeuSerProGlnLeuAspProLeuTrpArgIlyar	1097
Db	9744	GCACGACCGGCCGACCTTGACACCTGAGCCGCCACGACGCGCTGTGCGTGAAG	9803
QY	1097	gPro 1098	
Db	9804	ACCC 9807	
RESULT 11			
LOCUS	AF034576	3644 bp	linear
DEFINITION	Gallus gallus Janus tyrosine kinase (JAK) mRNA, complete cds.		
ACCESSION	AF034576		
VERSION	AF034576.1	GI:2645986	
KEYWORDS			
SOURCE	Gallus gallus.		
ORGANISM	Gallus gallus.		
REFERENCE			
AUTHORS	Archelos, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Archosauaria; Aves; Neognathae; Galliformes; Phasianidae;		
JOURNAL	Phasianidae; Gallus.		
MEDLINE	1 (bases 1 to 3644)		
PUBMED	Sofer, L., Kampa, D. and Burnside, J.		
AUTHORS	Molecular cloning of a chicken JAK homolog from activated T cells		
TITLE	Gene 215 (1), 29-36 (1998)		
JOURNAL	98332716		
Submitted (14-NOV-1997)	9666067		
Direct Submission	2 (bases 1 to 3644)		
	Sofer, L., Kampa, D. and Burnside, J.		
	Submitted (14-NOV-1997)		





```

source          1. .2914
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
BASE COUNT      566 a 907 c 809 g 632 t
ORIGIN
Alignment Scores:
Pred. No.:      2.85e-189      Length:      2914
Score:          3228.50        Matches:      614
Percent Similarity: 87.63%     Conservative: 45
Best Local Similarity: 81.65%  Mismatches:   90
Query Match:    53.09%        Indels:        3
DB:             Gaps:         3
US-09-397-967-16 (1-1099) x AX203084 (1-2914)
OY 350 ASPSERATGHSITRPHSCYLSGLVVALAIPROPRATGLeuGLUGLUALA 369
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 5 GACTCCAGCACTCTCTGCAAGAGAGTGGCAGCCGCGAGCTGCTGAGAGAGTGGCC 64
OY 370 ASPVALCYSHISGLYPROILETHRLAUSPHEALAILHISLVSLEULYALAGLY 389
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 65 GAGGAGTCCAGAGGCCCATCACTGTGGCATCAACAGCTCAAGACTGGGGCC 124
OY 390 SERLEUPROGLYTHRTYRILELEUARGARSPROGINASPTYRASPSPHELEU 409
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 125 TCACGTCCTGGCTCTATGTTCTCCGCGAGCCGCCAGACTTGGACAGCTTCTCTC 184
OY 410 THRALACYSVALGINTHPROLEUGLYPROASPTYRGLYSLEULIARGGLINASP 429
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 185 ACTGCTGTGTCACAAACCCCTTGGCTGTATTAAGGGCTGCTCATCCGCGAGC 244
OY 430 PROSERGLYALAPHSERLEUVALISLYLEUSERGINPROHISARGSERLEUARG 449
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 245 CCCACAGAACCTCTCTTGTGGTGGCTGCACCCAGCCCAAGCAGCATCTTGAGAGCTC 304
OY 450 LEUVALIACYSSTRIPASNSERGLYLEUARGVALIASPGLYALALALEUPLYRHSER 469
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 305 CTGGGCACTGCTGGATGGAGGGGCTGCACGTAGATGGGTGGAGTACCCTTCTCC 364
OY 470 CYSYALAPROARPROLYSGLYLSYSERASNLLEUVALIARGARGLYCYASN 489
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 365 TCGTGATCCCGAGACCCCAAGAAAAGTCCAACTGATGCTGCAGAGAGGTACAGC 424
OY 490 PROALIPROALIPROGLYSSERPRO---SERCYSYALALEUPLYRHSERPH 508
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 425 CCACCCACATCACTCTGTTCAAGCCCAATCCCAATACCAGCTGAGTACATTT 484
OY 509 HISTHRLIETPROTHASPSERLEUGLUTRPHISGLIASNLEUGLHISGLYSERPHET 528
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 485 CACAAAGATCCCTGTCAGAGCCGTGAGTGGCATGAGAACCTGGGCGCATGGTCTTCACC 544
OY 529 LYSILEPHEARGLYSERARGARGGLIUALIASPGLYGLUTHRHISASPSERGLIUAL 548
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 545 AATATTACCGGGGCTGTGCGCATGAGTGGTGGAGGGAGGCCGAAGACAGAGGTG 604
OY 549 LEULEUYSVALMETASPSERARGHISARGASNCYSMETGLUSERPHELEUGLUALA 568
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 605 CTCTCTAAGGTCTATGATGCGACAGACAAAGAACTGCATGAGTCACTTCGGAAGCAGCG 664
OY 569 SERLEUMETSERGINVALSERTYRPROHISLEUVALLEUENHISGLYALCYMETALA 588
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 665 AGCTTATGATGAGCCAAAGTGTGTCACGCGCATCTGCTGTCACGCGCTGTCAGGCT 724
OY 589 GLYASPSERTLEMETVALGINGLUPHEVALTYRLEUGLVALAILLESPMETYRLEUARG 608
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 725 GGAAGACAGCCCAATGGGAGGAAATTGTACACTGGGGGCCATAGACATGTATCTGGCA 784
OY 609 LYSARGGLYHISLEUVALISERALSERTPLYSLEUGLINALHRLYSGLINLEUALATYR 628
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 785 AAGCGTGGCCACCTGGTGGCCAGCCAGCTGGAAGCTGAGTGTCAAAAGCTGGGCTTAC 844
OY 629 ALALEUANSNTYRLEUGLUALSPRYSGLYLEUPROHISGLYSNVALISERLIAARGLYSVAL 848

```

```

DB 845 GCCCTTAATATCTGAGAGACAAAGCCCTGCCCATAGGAAAGTCTCTCCGGAAGGTG 904
    |||.....|||.....|||.....|||.....|||.....|||.....|||
OY 649 LEULEUALARGGLUGLYGLYASPLYASNPORPHLEILEYLSLEUSERSPROGLY 668
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 905 CTCCTGCTCGGAGGAGGCTGATGGAGACCCTGCTCATCAAGCTGAGTACCTGGG 964
OY 669 VALSERPROTHRALLEUSERLEUGLUMETLEUTHASPARGLIETPROTYVALAPRO 688
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 965 GTACGCCCTCTGTAAAGCTGGAGATGCTACCGAAGATCCCTCGGGTGGCCCC 1024
OY 689 GLUCYSLEUGLINALAGINTHRLCYLSLEUGLUALIASPLYSTYRPLYPHEGLYALA 708
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1025 GAGTGTCTCCGGAGGGCGAGACACTAGCTTGAAGCTGACAAAGTGGGCTTGGCGGCC 1084
OY 709 THRTHTRPLUVALPHEGINARGGLYPROALHISILETHSERLEUGLUPROALATYS 728
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1085 ACGTCTGGGAGAGTGTAACTGGCGTCACCATGCCCATAGTCCCTGGATCTGTAAAG 1144
OY 729 LYSLEUYSRPHETRYRGLIUAASPLINGLILEUPROALALEUYSRTPHRLIUEUALA 748
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1145 AACTCCAAATTTATATGAGACCGGACAGCTGCGGCCCCCAAGTGGACAGAGTGGCC 1204
OY 749 GLYLEULETHRGLINCYSMETALATYRASPPOGLYARGARPROSERPHEATGALAIL 768
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1205 CTGCTGATTCACAGCTCATGAGCTTATGAGCCGTCAGAGGCCCTCTCCGAGCCGTC 1264
OY 769 LEUARGASPLEUANSGLYLEUILETHRSERASPTYRGLIUELEUSERASPTROTHPRO 788
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1265 ATTGCTAATCTCAATACCTCATCTCTTCAAGACATGAGTCTCTCCAGACCCACACT 1324
OY 789 GLYLEPROSERPROARGASPLILEUCYSVALAIALIYALAGLILEUPLYRALSGLIN 808
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1325 GGTGCCCCGAGCACTCGTGAATGGGCTGTGG---AATGGTCCCAAGCTCTATGCTGCCAA 1381
OY 809 ASPPROALATIEPHEGLUGLUTRPHISLEUPLYRHSERLEUENHISGLYSGLYASN 828
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1382 GACCCACAGATCTTGAGAGAGACACCTCAAGTACATCTCACAGCTGGGCAAGGGCAAC 1441
OY 829 PHEGLYSERVALISLYLEUCYSARGTYRASPPOLEUGLYASNPANTRGLYPROLEUVAL 848
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1442 TTGGCAGACGAGGAGCTGTGCGCTATGATGACCCGCTACCCCAATACAGGTGCCCTGGTG 1501
OY 849 ALAYALLYSGLINLEUGLHISERVALPROASPLINGLINARGASRPHAGLITILE 868
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1502 GCCGGAACAGCTGCACACAGCGGGCCACAGACCAAGAGACTTTCAGCGGGAATT 1561
OY 869 GINILEUYSALALEUENHISERASRPHLEVALIYSTYRARGLYVALISERTYGLY 888
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1562 CAGATCCCTCAAGACACTGCACAGTATTTCAATGCAAGTATCGTGTGTCAGCTATGCGC 1621
OY 889 PROGLIYARGLINSERLEUARGLEUVALMETGLUTYRLEUPROSERGLYCYLSLEUARGASP 908
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1622 CCGGGCCGCGCAGAGACTCGGCTGTCATGAGATGACCTGCGGAGGGTGGCTTGGCGGAC 1661
OY 909 LEUENHISARGHISARGGLY---LEUHSITHRASPARGLEULEUENHISALATRPLIN 927
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1682 TTCTCTGACAGCGGACACCGCGCGCTGATGACAGCGCCCTCTCTTATCTCTCCGAG 1741
OY 928 ILECYLSGLYMETGLUTYRLEUGLYALARGARGCYVALIHISARGASPLEUVALA 947
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1742 AACTGCAAGAGGATGAGTACTGCTGGGCTCCGCGCTCTGTCAGCCGAGCTGGCGGCC 1801
OY 948 ARGASNILLEUVALIGLUSERGLIUALHISVALIYSILEALIASRPHGLYLEUALALYS 967
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1802 CGAAACATCTCTGTGAAGAGCGAGGACACAGTCAAGATGCTGACTTGGCTTACCTAAG 1861
OY 968 LEULEUPROLEUGLYLYASPTYRGLYVALIARGLUPROGLYGLINSERPROLIPHE 987
    |||.....|||.....|||.....|||.....|||.....|||.....|||
DB 1862 CTGCTGCCCTTGACAAAGACTACGATGCGTCCGAGCGCAGGCCAAGCCCATTTTTC 1921
OY 988 TRPYRALAPROGLUSERLEUSERASPNLIEPHESEARGLINSEVALTYRPRER 1007
    |||.....|||.....|||.....|||.....|||.....|||.....|||

```







QY 140 LnhlsrSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGlnG 160  
 |||||  
 Db 1460 AGCAGCGCATGTAGTGTAGTGGGGCCCTCCCTGGGGCTAGTCTCAAGAGCAGG 1519  
 QY 160 LysGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnAlaGln 180  
 |||||  
 Db 1520 GTGAGTCTCTCAAGCTGCGCCCTGTTGGACCTGCGCCGATGGCGCAGACAGCCCAAGC 1579  
 QY 180 rPrgLysGluLeuLeuLysThrVal----- 188  
 |||||  
 Db 1580 GCGCGGGAGACCTCTGTAAGACTCTCAGGTAGAGACCACCAAGCTGTGGGACGGGCTCT 1639  
 QY 188 ----- 188  
 Db 1640 GTTTGGAGTAGCAACGTGGGCTCATCGGGGGTTGGCGGGCTCCACCATGAGATTTC 1699  
 QY 188 ----- 188  
 Db 1700 TCCGCAAGCTTTCAGGGTGTCTCTATGACCCAGGGGCTCCACGAACCCAGGCTCTCC 1759  
 QY 188 ----- 188  
 Db 1760 ACCCCAGGCTCAGCTGACAGACACCTGACCTCCCCAGCTAAAGGCTGTGGGCTC 1819  
 QY 188 ----- 188  
 Db 1820 CCTGTCCGACCTGTGGCGGCAATGGCCCTCCCTACTGTAGGTCCGGTCTCATACCT 1879  
 QY 188 ----- 188  
 Db 1880 GACCCGTAATGAGAGTCTGTGTGTGCTGCTGCTGCCCACTAGGGCGGCACCCAGCCCT 1939  
 QY 188 ----- 188  
 Db 1940 GGGCTAAAGCTGGTGTGTGTGTACCCCGGGGACCCCTCCGACGCTGAGAGCG 1999  
 QY 189 ----- SerTyrLysAlaCysLeuProProSerLeuAlaGlyAspV 201  
 |||||  
 Db 2000 GGTCCCTCCCTCCCAACCCCTGACAGCTACAGGCTGTCTACCCCAACCCCTGCGGAC 2059  
 QY 201 aLlLeGlnGlyLnaPheValThrArgArgArgLleaArgLghThrValAlaLeuAlaL 221  
 |||||  
 Db 2060 TGTATCCAGGGGCTGACCTCTGTGACCGGAGGCTATTCGGAGACGGTGGCCAGAGCCC 2119  
 QY 221 euleuPrgCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLeu 240  
 |||||  
 Db 2120 TGCC-GCGGTGGCCCGCTGCGCAGGACCGGCACTGCTCATGCGCAAGTACATCATG 2178  
 QY 241 AspLeuGluArgLeuHisProAlaAlaThrThrPheArgValGlyLeuProGly 260  
 |||||  
 Db 2179 GACCTGAGCGGCTGTGATCCACCGGGGGCGCGGACCTTCCACGTGGGCTCCCTGGG 2238  
 QY 261 Ala-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyLe-Pr 277  
 |||||  
 Db 2239 GCGCTTGTGGCCACGACGCGCTGGGGCTTCCGCTGTGCTGTGACGGCGCATCCGC 2298  
 QY 277 0----- 277  
 Db 2299 TGGACCCAGGAGAACAGGGGTGAGACGGCTTCCCCCTGCGGGGCCAATGGGGGGG 2358  
 QY 277 ----- 277  
 Db 2359 GTTTGGGAGGGCGGAGATTGTAGGGGATGTGGGGCGGAGCCAAACGAAGACTTG 2418  
 QY 277 ----- 277  
 Db 2419 GGGAACTGGGCGAGGCTTAATGAGGGCGGGCTTAGTGAAGGAGAGACTGCGGAAATG 2478  
 QY 277 ----- 277  
 Db 2479 GGAGGGCAAACTGAGTAGAGGTTGGGCTGAGCGACCCGGGAATGAATAATGGAGGGGT 2558  
 QY 277 ----- 277

Db 2539 TGAGAGTGGGTGGTGGGGCTGGAAGCAGCAAGCTAAGAGCAGAGTAGGCTCCG 2598  
 QY 277 ----- 277  
 Db 2599 AGAGCTGGGAGGCTGGCAACCGGGGCAAGGATCAAGGGGGGTCGCGCAGAGAAAG 2658  
 QY 278 -----TrpSerSerAsnAspLysPhe----- 285  
 |||||  
 Db 2659 AGCGGGGATTTGGCAGAGCGGGTGGGAGATGTTGGGGAATGGGTAAGGAGGTGGAGA 2718  
 QY 285 ----- 285  
 Db 2719 GGTGGAGTTGATAGAGGGGCGGGCTTGGAAAGGTTGAATGCAAAAGGATAGGAG 2778  
 QY 285 ----- 285  
 Db 2779 TGGATGTGTGGCTTGGGGGTGGTGCATGGGCGTGGTTGGCGGACCCGCTGGCCC 2838  
 QY 286 -----GlnThrPheCysAspPheProGluIleVal 295  
 |||||  
 Db 2839 CACTAGGTTAANCCCAACACCCAGCTCTCCAGCCCTTCTGCGACTTCCAGAAATCGT 2898  
 QY 295 LAspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGlnHisArgLeuValTh 315  
 |||||  
 Db 2899 AGACATTTAGCATCAAGCAGAGCCCGCGCTTGGCCCGGAGAGCACCCGCTGTGTAC 2958  
 QY 315 rValThrArgMetAspGlyHisIleLeu----- 324  
 |||||  
 Db 2959 TGTTCACGAGACAGACAAACAGATTGTTAGTGTGTGACAGATTCCCTCCCTTCAGCCT 3018  
 QY 324 ----- 324  
 Db 3019 TACCCCGAAGGCGGGACCGGACCCCTCGGTTTCACGTGGGCTGTGACGCTTGTCCCTCG 3078  
 QY 325 --GluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyT 344  
 |||||  
 Db 3079 CAGAGCGCGAGTTCACAGGGCTGCGCGAGGCTGTGTGCTGTGGCGCTGTGACAGGCT 3138  
 QY 344 YrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgL 364  
 |||||  
 Db 3139 ACTTCGCGGCTGACCAAGCGGCTCCAGCAGCTCTCTGTGCAAGAGGTGGCACCGCGGAGC 3198  
 QY 364 euleuGluGluAlaAlaPheValCysHisGlyProIleThrLeuAspPheAlaIleHisL 384  
 |||||  
 Db 3199 TGCTGAGAGAGTGGCGGAGAGTGCACGSCCCCATCTCTGAGCTTGGCATACACA 3258  
 QY 384 YsLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspT 404  
 |||||  
 Db 3259 AGCTCAAGACHTGGGGGCTCACGCTCTGCTCTATGTTCTCCGCGCAGCCCGCAGGACT 3318  
 QY 404 YrAspSerPheLeuLeuThrAlaCysValGln-Thr----- 415  
 |||||  
 Db 3319 TTGACAGCTTCTCTCTACTGTCTGTGTGTCAGAGGTCTACTGTAGGGTGGGTACTGG 3378  
 QY 415 ----- 415  
 Db 3379 AGGCTGCTGAGACAGAGTTTACCTCTGAATTGAGATTACTCAGAACTACTCAGCAT 3438  
 QY 415 ----- 415  
 Db 3439 TTGGTTCCTGGAACAAATAGAGGAAACGACACCCCGGTTAAGGGGAAACAGCGTATTCT 3498  
 QY 415 ----- 415  
 Db 3499 CTGCTTAATTAAGTATGATTGTTGTTGCTTTCAGAGATTAAGACTGAAGCTGGGGCG 3558  
 QY 415 ----- 415  
 Db 3559 TGTGCGGTCGCCGTTTATCCAGCCCTTGGGAGGCTGACGTAAAGCATATCCCGAGATC 3618  
 QY 415 ----- 415

Db	3619	GGGATATCCACCCCATCCTGGTCCAAATGTTAAACCCTGTCCACTGAATATGTAA	3678
Oy	415	-----	415
Db	3679	CTCAGCTGGGCATGGTAATGGGCATCTGTAAATCCACTACTCGGGAAAGCTCAAGGCAT	3738
Oy	415	-----	415
Db	3739	GTTTTCTAAGTTTTTAAAAAAGGGGGGGGGCTTTTTTGGGGTTTTTAGGGAAAAAAA	3798
Oy	415	-----	415
Db	3799	AAGGGGTGGGGGAAAAAAGGGGTAAAAACCTTTTTTTAGTCTTGAGAGCTCCTGGCAAC	3858
Oy	415	-----	415
Db	3859	CTCGTGGAGATGGGGGGCTGCACGTAGATGGGTGGCAGTACCTCACTCTCTGCTGAC	3918
Oy	415	-----	415
Db	3919	CCCAAAATTTGGGGGGGTGGGGTTGGCCAGTTAAAGTTTAAAGTTTGAAGTTGNAACCTTCGGG	3978
Oy	415	-----	415
Db	3979	GCCCAAAAGGGGGAACCCCTTTTTTTTAAAAAATTAAAAATTTTGGGGGGGTAAAG	4038
Oy	415	-----	415
Db	4039	GGGATTTTTTTTCCCCCAAAATTTGGGGGGGGGGCGGGCATTTTTTAAACCCGGGG	4098
Oy	415	-----	415
Db	4099	GGGGGGGGTTCCAGGGGCCAGAGATAAACCTTTAAATTTCCCGGGGGAAGGCCCA	4158
Oy	415	-----	415
Db	4159	AATTTCTTTTCCAAAAAAGAAAAAAGAAAAAGAGAGACTGAGAAAGAG	4218
Oy	416	-----	419
Db	4219	AGTGCTGTGCTCAGTCCCACTCAGGGGCCACTCTTCTTTCAGAAACCCCTTGGTCT	4278
Oy	420	ASPTTTLysGLysLeuLLeaTArgLInaSPProSerGlyAlaPheSerLeuValGlyLeu	439
Db	4279	GATTTAAAGGCTCTCCATCGGGGCCAGCCCAAGAGAACCTTCTCTGTGGTGCCTC	4338
Oy	440	SerGlnProHisArgSerLeuAArgLLeuLeuAlaIaIaCysTTPasSerGlyLeuArg	459
Db	4339	AGCCCAACCCCAACACACTCTTCGAGAGCTCTGGCAACCTGTGGATGGGGGGCTGCAC	4398
Oy	460	ValAspGlyAlaIaIaLeuTyrLeuThrSerCysAlaProArg-ProLys-----	476
Db	4399	GTAGATGGGGTGGCAGTGAACCTCACTCTCTGCTGTATGCCAGAACCAAGGTGAGGCC	4458
Oy	476	-----	476
Db	4459	CCTGTGCCCTCGAATGAGTGGCTGATCTGGACCCCTGTTTCTATGTCTGATTAACAG	4518
Oy	476	-----	476
Db	4519	CTCTGTGTGATGGCAAGTGTGTGAAGAACTGCAGATCATGTGGGTTAGAGAAAGAAAGG	4578
Oy	476	-----	476
Db	4579	CGATTTGTTGCTCAGAGAGTTAGACCATCTAAACCCTAAGGGCTGGAATTGCCAGGTA	4638
Oy	476	-----	476
Db	4639	CAGTAGATGTAGGCCATGTCACTGGACCTGTGNTTATTCACAGATCTCAGTTCTTTC	4698
Oy	476	-----	476
Db	4699	TTCTTGGTGACCTTCACTCCTTTGGCACAACCTCCCCCTGATAGGGAGCAAAATTACTT	4758

OY	476	----	476
Db	4759	GAGTATGTGGAGCCCCCTTCTCTGAGTCTTCCAGCAATTCCTCGGTAAAGNAACCGAT	4818
OY	476	----	476
Db	4819	TCCCTTGGGCTCGCTCCACAAAGTGTCTAAGCTTTGGCTTCAATTAATACGGCCCCCTGA	4878
OY	476	----	476
Db	4879	AAGGGGTTTCCTTCCCTTCCCTTGGGGGGGTTTCACCCCCCAACAAATTCCTCATTG	4938
OY	476	----	476
Db	4939	GGGGGGGCTTCTTTGATTTTTCCTTCCCAAAATAAAACCATTGTATAAGTTGGC	4998
OY	476	----	476
Db	4999	GCCACAGAAAGATGGAGAGGTGGAGAGAGAGAGATGCGGGGGTCTTCTCACCTCC	5058
OY	476	----	476
Db	5059	AATAATCCCAATTTTGTGTTCACAGACCTGCTTAGGGGAAACCTGGCACAGGTCAAC	5118
OY	476	----	476
Db	5119	CTTTGGCCCTGGAGTTTGCATTAGTGCACACAGACAGTACATGATTTTTTGTATACGGG	5178
OY	476	----	476
Db	5179	ACAGAGTTCCTACTATCATCTATGGCTTTTGTGTGATTGCTTTGTGTAGACAG	5238
OY	476	----	476
Db	5239	AGTCTCTTGTGCACCCAGGTGAGTGCAGCGGCGCATTTGTATTTACTGCACCTC	5298
OY	476	----	476
Db	5299	CGCTTCCAGATTCAGCAATTTCTCTGCTTCCTCCAGTTCAGTGGGATTAAACG	5358
OY	476	----	476
Db	5359	GCATGCACACACGCGCGCTAATTTGTATTTAGTAGACGCGGTTCTTCATG	5418
OY	476	----	476
Db	5419	TTGTCAGCGTGTCTGCACCTCCGACCTCAGATGATCTGCCGCTTACGCTTCCAA	5478
OY	476	----	476
Db	5479	GTGTGTGATTCACAGGCTGAGCACCGCGCGCTAAATTTTTTGTCAAGACGAG	5538
OY	476	----	476
Db	5539	GTCTGCTATGTGCTAGCTGTTTTGAACGTGTGGCTTAACACATCTCTCTCTCA	5598
OY	476	----	476
Db	5599	GCTCCAAAGTGTGAGACAGGTTTCCCGCTTCTTTAAAMACCAACCCACCGG	5658
OY	476	----	476
Db	5659	CCCCAATGTGTGAACCTTTATTAAACCCAAACTTTGCGGGGAGAGGGGGGTT	5718
OY	476	----	476
Db	5719	TAGTCCCCCGGGGTCTTTTGGGGTTTGGCAACCCCGCCCTTAAAAAAGGGGCAA	5778
OY	476	----	476
Db	5779	ACCCTTTTTCCGCCATAAAAAAGAAAAAGAGTGTGGAGGGTTGAATAATTCGTG	5838

```
QY 476 ----- 476
Db 5839 TTTTCCCAATAATCCGTGGGAGAGGCGGGGGGCTTACTTTGACCCCGGGGTCGGGT 5898
QY 476 ----- 476
Db 5899 GTGGTTCCTCCGATTAACACAGAAATTTCCCGGGGAAACCGGGGAATTTTTT 5958
QY 476 ----- 476
Db 5959 TTTTAAAGAACAAAAAATTAGCGGGGACGGGTGGCACACACCTGTATGCCAGC 6018
QY 476 ----- 476
Db 6019 TATTGGGAGCGGAGGAGGAAATCGCTTGAACCTGGAGGTGGAGTTGCAGTAGC 6078
QY 476 ----- 476
Db 6079 TGAGATCGACCATTTGCCACCACCGCTTGATGACAGAGTAGATCCGCTTCAGCAGCAG 6138
QY 476 ----- 476
Db 6139 CAGCAACAGCAAAACAAACAAACAAAGCCATGTGCCCTGAAGTCTTCATCTCAG 6198
QY 476 ----- 476
Db 6199 GGTGGGTTCTAGAGGGTACTCAACTAAGCATGAGTAGCTAACCTTGGGGACTTT 6258
QY 477 ----- 477
Db 6259 TCACCTCTGATTTCTGGTTTTTTCCTCCCATCTCTCCCTCCATAGAAAATCCACCTGAT 6318
QY 482 eValValArgrglLycysnProAlaProAlaProGlyCysSerPro---SerCysCys 501
Db 6319 CGTGCTCAGAGAGCTCACAGCCACCATCATCTCTTGGTTCAGCCCAATCCCAATA 6378
QY 501 salaleuThrlInleuSerPheHsThrIleProThrAspSerleuGluTrp----- 518
Db 6379 CCACCTGAGTCAGATGACATTTTCCACAGATCCCTGCTGACACCTGGAGTGGTAAGAG 6438
QY 518 ----- 518
Db 6439 CCCTGGAAATGAGGCGATACCTGAGTGGGTCCAGAGACTCAGATGCTGGCTCAG 6498
QY 518 ----- 518
Db 6499 GCATATGCTATATTTTACCTTGCCTTCGGTTTTTCCATCTGTAAATGGGGCCAGCAGCT 6558
QY 518 ----- 518
Db 6559 ATGTCTGGCTTGGGCTGGGATCCTGCAGAAACCCCTCACTGGCCTTCTGTGTCCTCC 6618
QY 519 ----- 519
Db 6619 TCACCATTCAGCATGAGAAACCTGGGCCATGGGCTTTCACCAAGATTACCGGGCTGTGC 6678
QY 535 rGrArGlValAlaAspGlyIuThrHsAspSerGluValLeuLeuLysValIleMetAspS 555
Db 6679 GCCATAGCTGGTGGATGGGAGGCGGCAAGACAGAAAGTGTGCTGAAGGTCATGATG 6738
QY 555 eFrlHsIAsrAnCysMetGluSer----- 563
Db 6739 CCAAGCACAGAACTGCATGAGGT- GAGAGGCAATGTGACAGACACTTTGGAGTGGG 6797
QY 563 ----- 563
Db 6798 GGCTGGCTTGAGCGGGGCTCGTGATGCAGAGACATTTACAACACACAGGAGACTGGGC 6857
QY 563 ----- 563
Db 6858 GTGGTGGCTGATGCTGCACATCATAGCGGGAGCTCGTCTTGACAAAAAATTTTAGGCC 6917
QY 563 ----- 563

Db 6918 GGGCGCGGTGCTCACGCTGTAATCCACAGACTTTGGGAAGCTGAAGCGGAGAAATCA 6977
QY 563 ----- 563
Db 6978 TCGAGGTGGGAGTCTCGAGACCAACCTGCCCAACATGAGAGAAACCGGTCTACTAA 7037
QY 563 ----- 563
Db 7038 AAAACCAAAATTTACCGGGCGTGGTGTGCATGCTGTATATCCAGCTACTCGGGAGTC 7097
QY 563 ----- 563
Db 7098 TGAGCAGAGAAATCTCTTGAATCTGGAAACGAGAGTTCCAGTGATCAAGATCGCGCC 7157
QY 563 ----- 563
Db 7158 ACTGCATCCAGCCTGGGTGACAGAGGAGACTGTCTCAAAAAACAANCAANCAANA 7217
QY 563 ----- 563
Db 7218 CAAAAACCATAGATGATAGTGGAACTTGTGCCGTATCAGAAATCATGTAGTGTCT 7277
QY 563 ----- 563
Db 7278 GTGTGCATTAATGGCAGACTCCAGGGCCAAAGTGACCTGTGGCCAGGTTCCTTAAG 7337
QY 563 ----- 563
Db 7338 GCAGGTCTGTGAGCACAAATTTGGGATTATTGAGTGGAAACCCACCATCTTCTC 7397
QY 564 ----- 564
Db 7398 TCCCTTCCACCTTCCAGTCATTCCTGTGGAAGCGCAGCCTTGATGAGCAAGTCTGT 7457
QY 576 yPProHsLeuValLeuLeuHsGlyValLysMetAlaGlyAsp----- 590
Db 7458 ACCGCGATCTGTGTGCTGCCAGCGGTGTGCATGTGCTGGAGACAGGTGACTGCTCTAG 7517
QY 590 ----- 590
Db 7518 GGGGTGTTAAAGGGAGAGGAGGAGAAAGAGCGGAGGGGAAAGAGCTAGGGGCA 7577
QY 590 ----- 590
Db 7578 GAGAGAGAGAGAAAGCCGGGAGAGGCGCAACGGAGACAAAGAGAGTTAGCGAGA 7637
QY 590 ----- 590
Db 7638 GGATAAAAAACTGGCTGACGTTTATAAAGATAGGGGGGAGAAAGCTTCCTCGAAGC 7697
QY 590 ----- 590
Db 7698 AAATTATATAGGCCCGGAGAGAGACTTGTGGCTTGGGGGGGCTGAGAGAGAACCC 7757
QY 590 ----- 590
Db 7758 GCCCTTTAAACAAACTAGAAAGACCTGGGGGATATACCCCGCTTATTTATAACCTGG 7817
QY 590 ----- 590
Db 7818 AGAGAACTCAAGAAAAAGAGATCCCTGGATGCCCCAGATATATTTCGGAGTGGCC 7877
QY 591 ----- 591
Db 7878 AAAGAACTAAACTTGGAAATCTGAGCAATATAGGCTGAGCACACATGGTGGAGAAATTTG 7937
QY 598 aLtyrLeuGlyAlaIleAspMetTyrLeuArgLysArgLysIleValSerAlaSerT 618
Db 7938 TACACCTGGGGGCGCATGACATGTATCTCGAAACGTGGCCACCTGTGGCCAGGACAGT 7997
QY 618 rPlysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyr----- 632
|||||
```

```
Db 7998 GGAAGCTGCAGGTGTCACAAACAGCTGGCCCTACGCCCTCACTATCTGTAGTCTCTC 8057
Qy 632 ----- 632
Db 8058 CCCCAGGAGATTTGTTTGGGGAGGGGGGGTTAAAAAACCCAAACCCCTTAAT 8117
Qy 632 ----- 632
Db 8118 TTTGGGGGGGGTGGTGGCTGTTCCCTCCATTCCTCCGGGAAGTTTTTTTTTTT 8177
Qy 632 ----- 632
Db 8178 AAAGGGAAGAAGATGTTTTTTTTTTTAAAGATTGTGAAGTTGGGGGGGG 8237
Qy 632 ----- 632
Db 8238 CATTTTACTTTCTTTTGGAAAAATGGAAGAAAAATTTTTTTTTTTTNGGGT 8297
Qy 632 ----- 632
Db 8298 TTGGGGAAGGGGGTGGCCGGAGAAAGTTTCTTTTGGAGAAAAAGGACATGCC 8357
Qy 632 ----- 632
Db 8358 CTTACCTCGAGTGGGTTTTTTCATGAAGAAGAGAGAGATTCGGTTTGGGGA 8417
Qy 632 ----- 632
Db 8418 ACCAAAGCAAGCCGCGCGGAGAAAGAGATTTCCCTGGCCCTATCAGCATT 8477
Qy 632 ----- 632
Db 8478 TCCGTTGCCAGAGACAGGAAAAAGTGGAGAGGAATTAATTGAAGCTTGAGCAG 8537
Qy 632 ----- 632
Db 8538 GGGCTGGAATCCAGGTGAGAGTTAATTTTTTCAAGGATTTGGGTTAACAG 8597
Qy 632 ----- 632
Db 8598 ATGGGTGTGTTTGAAGGGGAGAGACATGATGATTTTGGGGGTGAGTAGAG 8657
Qy 632 ----- 632
Db 8658 GGGCAAGACAGATATCAGATTTCAAGGACATGTTGAGCAGTACCAA 8717
Qy 632 ----- 632
Db 8718 GTGGGTTTGAAGATATAGAGTTTGGCAACAGACTCTTCAATTTATCAACCTCC 8777
Qy 633 ----- 633
Db 8778 CCGGCATTTCTGTGTGTTGGCCCCCTTAGAGAGACAAAGCCTGCCCATGTGTC 8837
Qy 644 SerAlaTrpValLeuLeuAlaArgLysIleValAspLysProPheIleVal 643
Db 8838 TCTGCCGGAAGTGTCTCGGCTCGGAGGGGCTGATGGAGCCGCCCTTCATCAAG 8897
Qy 664 LeuSerAspProGlyValSerProThrValLeuSerLeuVal 677
Db 8898 CTGAGTACCTCGGGTACGCCCGCTGTGTTAAGCTTGAGGTAACTCTCGAGGTG 8957
Qy 677 ----- 677
Db 8958 AGAAGGAGGGCTGACAGAGGCAAGAGATCCCTGATCCCACTTCATTCCTCA 9017
Qy 678 MetLeuThrAspArgLysProTrpValAlaProGluCysLeuGlnIleValIle 697
Db 9018 GATGCTACCGACAGATCCCTGCTGGTGGCCCGAGAGTCTCGGAGGCGGACACT 9077
Qy 697 uCysLeuGlnIleAspLysTrpGlyPheGlyAlaIleThrTrpValPheGlnArg 717
Db 9078 TAGCTTGAAGCTGACAGAGTGGGCTTCGGCCACGGCTTGGGAAGTGTAGTGGGT 9137

Qy 717 yProAlaHisIleThrSerLeuGlnProAlaLys 728
Db 9138 CACATGCCCATCAGTACGCTGATCTGATAGGTGACAGCCCTCACCGGCATGGT 9197
Qy 728 ----- 728
Db 9198 CTCCGAACCCCACTTTGACAGCCGGGCACTGACATCCAGTCTGGGAGATTGGGGT 9257
Qy 729 ----- 729
Db 9258 GGTCTATTGGGTTGGGATTAACGACTGCTCTCACCCTCAGAACTCCATTTTATG 9317
Qy 734 LysPheGlnGlyGlnLeuProAlaLeuLysTrpPheGlnLeuAlaGlyLeuIleThrGln 754
Db 9318 AGACCGGACAGACGCTGCCGGCCCAAGTGCAGAGAGTGGCCCTGCTGATTCAACAGT 9377
Qy 754 yMetaIaTrpAspProGlyArgArgProSerPheArgAlaIleLeuAlaGlyLeuAsn 774
Db 9378 GCATGGCTATGAGCCGCTCAGAGGCTCCTCCGAGCCGCTCATTCGACTCAATA 9437
Qy 774 LysLeuIleThrSer 778
Db 9438 GCTCATCTCTCAGGTGCCCTGGAGCGGTGGGTGGGAGGCTGTGATCATAT 9497
Qy 778 ----- 778
Db 9498 TGGCCAGTGAAGAGCGTGGTTGCACTAGGCCACGCGCTGTGTGTGAGTT 9557
Qy 778 ----- 778
Db 9558 GGAGGGGTGATGACTATGACAGACAGTGTAGCTTCAACAGTACTGTAAGACAGTGC 9617
Qy 778 ----- 778
Db 9618 CCTCAGTTCAAGTTCAACATGATGATTTATTTATTTATTTATTTTGAAGACAG 9677
Qy 778 ----- 778
Db 9678 TCTTGTCTGTGGCTAGAGCTGAGTGCAGTGCACAATTCGGCTCAGTGAACCTCA 9737
Qy 778 ----- 778
Db 9738 CCTCTTAACCTCAGTCAAGGATTCCTCCTCAGACTCCTGAGTAGTGAGATT 9797
Qy 778 ----- 778
Db 9798 ACAGGTGTGCACACGCGCCAGCTAATTTTGTGTTTTTGGAGAGATTTGGGGT 9857
Qy 778 ----- 778
Db 9858 TTCAACCATATTTGGCTAGGGTGTCTTGAACCTTGACCCCAAGTGATCCGCCAC 9917
Qy 778 ----- 778
Db 9918 CTCTACCTCTAAGAGTGTGGATTCAATTTCTCTGACTTGCTTGTGTAATAT 9977
Qy 778 ----- 778
Db 9978 TGCCAATTAGTCGATATAGGTTTGCACAGTAGACTCATGATTCATTCGTGAG 10037
Qy 778 ----- 778
Db 10038 ACTTGTAAACAAGATTTCAGAGTGTCAAGAGAGAACCGGATTTGTTATAAATAA 10097
Qy 778 ----- 778
Db 10098 CCTGACATATTAATAGAAATCAACATAAATAAGACTCAGTATAATACAGGAGAACAC 10157
Qy 778 ----- 778
Db 10158 CCAATAAGATTTAAGTTGCCCTTAAGAAAAAAGAAATGGCGTATGTAATAATTTTA 10217
```

```

Qy 778 ----- 778
Db 10218 AAGTCCAGCCAAACATATATAGCGTATAGATTCTGTGTAACAGTGAACATAGCAAAA 10217
Qy 778 ----- 778
Db 10278 TCCACACAAAGAACCCAAAAAGCAGCTACAGATCATTTCAAGAGTGTGAATGTGANT 10337
Qy 778 ----- 778
Db 10338 CTTCCTCAATTTTGGGCTGTCCCTCCTTCACAAATTCATCTTTTCAAAAGGCAATTTTG 10397
Qy 778 ----- 778
Db 10398 AAGATATTAGGTAGAGAGAGAGTGTGTCTATCTGTCTCAACTATGAGGCAATGCA 10457
Qy 778 ----- 778
Db 10458 GAGTGACGTGAATATATAGCATTTTGTATGAGTCAGCGCCATATAGTGTATAG 10517
Qy 778 ----- 778
Db 10518 ACCTTACTATATAGAACCGAGAAATATGATGTCTCCTCAATTTATGCCCCCATCC 10577
Qy 778 ----- 778
Db 10578 TTTGTAAACCAATTTGTGGGAAGTTGGCAGCGGTTTTCACCTGCACGAGGAGGCGC 10637
Qy 778 ----- 778
Db 10638 TTCTCTTCCCTGGAACAGAGTGGGAAGACACTTTTTCACACTCTCCACAAGGT 10697
Qy 778 ----- 778
Db 10698 ACCCTTTTGGCTTGCATATGCGCCCTGCCATATACAGAGAGGTCATATGCAATGGA 10757
Qy 778 ----- 778
Db 10758 GGTTCACACCAAGTCAAACTCAGAGAGTGGACCCAGATGAGAGCGCTCTTACCAC 10817
Qy 779 ----- 792
Db 10818 GCCCATGCCCCACCCAGCACTATGAGCTCTCTCAACACCCACACCTGGTCCCTG 10877
Qy 792 eRPProAargpLuleuCyValaLagLValaGlnLeuTyraLacysGlnaPProAaI 812
Db 10878 CACCTGTGATGTGTGTG---AATGTGCCAGCTCTATGCTGTGCAAGACCCACGA 10934
Qy 812 lPheGlnGlnaArgHsLeuLySTyrTLeSerLeuLeuLyS----- 826
Db 10935 TCTTCAGAGAGACACCTCAAGTACATCTCACAGCTGGGCA--GTAAGGTGGGCAAGG 10993
Qy 826 ----- 826
Db 10994 CCAGGTGGGTTGAGAGGCGAGCGACATCCAGTACCTGACATCAGTCCGCTATCC 11053
Qy 827 -----GlysnrheGlyserValGlnLeuCyAlaGlyTyrAspProLeuGlyAspAnthrG 845
Db 11054 TCAGAGCAACTTTTGGCAGCGGTGAGCTGTGCCCTATGACCCGCTAGGCCACATACAG 11113
Qy 845 lYProLeuValAlaValLysGlnLeuGlnHsSerValPProAspGlnaPProAspPheG 865
Db 11114 GTGCCCTGTGGCCCTGAAGACCTGCAGCACAGCGGCGGACAGCACAGGAGGACTTTC 11173
Qy 865 lNArgGlnGlnLeuLysAlaLeuHsSerAspPheLLeValLysTyrArgLys 885
Db 11174 AGCGGGAGATTCAATCTCTCAAGCAGCTGCACAGTATTCATTTGTAAGTATCGTGTG 11233
Qy 885 alserTyrGly----- 888
Db 11234 TCAGTATGGCGCGGTGAGCGCACTCCCGGATGATGAACAAGACGTATGGTCTTT 11293
Qy 888 ----- 888

```

---

```

Db 11294 TCAAGTGCACATTTTACCCTCCTGCCAGGCACTTTAGTATAGCTGGGAACGTATTA 11353
Qy 888 ----- 888
Db 11354 TTGATCCAGATTCACATGGGTAGCAGTTTGAATCTGACTACCCCTTACCAGTTGTG 11413
Qy 888 ----- 888
Db 11414 AAACCTTAACCTTTCTGAGCCCTCAGTTCTCCAGCTGAAGAAATGGCCATGTGATATT 11473
Qy 888 ----- 888
Db 11474 ACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11533
Qy 888 ----- 888
Db 11534 AGGCTGGAAGCAGTGGCTCAATCTCGGCTCACCTGCAACCTCTGCTTCCAAATTAAGT 11593
Qy 888 ----- 888
Db 11594 GATTCTCCTGCCCTCAGCCTCCATATATAACTGGGATTAACAGACGTTCTCAGATTAAC 11653
Qy 888 ----- 888
Db 11654 TGGGATACAGGAGGCGTCCACACACCGGCTAAATTTTGTGTGTGTGTGTGTGTGTGT 11713
Qy 888 ----- 888
Db 11714 TTTCAACATGTTGGCCAGCTGCTGTGTAACCTCTGACCTCAACTGAGGCTTCCAAAGTG 11773
Qy 888 ----- 888
Db 11774 CTGGGATTACAGGCGTGAGCCACTGACCTGCGCGTGATATATAACACATCATGTATG 11833
Qy 888 ----- 888
Db 11834 TTGTAGAAGTACTCAAGAGATATTACCAATGTGTATGTGTGCGGTCCCATATTAACAG 11893
Qy 888 ----- 888
Db 11894 TCACAAACTGAGTGCAGAGGAGGACACAAAGTCCCACTGTGAAGGGGGAAGATGGG 11953
Qy 889 -----ProGlyArgGlnSerL 894
Db 11954 GGGACGAGCAGGCGTGGCCCTCTGTACAGATTCCTGCTTCCAGCGCCGCAAG--C 12012
Qy 894 eUArgLeuValMetGluTyrLeuProSerGlyCysLeuArGAspLeuGlnaPProHsA 914
Db 12013 TGGGGGTGTCATGAGTACGTGCCAGGCGCTGCTTGGCGCACTTCTCTGCGCGGCAAC 12072
Qy 914 rGly---LeuHsThrAspArgLeuLeuLeuPheAlaTrpGlnLeuCysLys----- 930
Db 12073 GCGCGGCGCTCGATGCGCACCGCCGCTCTCTATTCCTCCGAGATCTGCAAGGTGAGAG 12132
Qy 930 ----- 930
Db 12133 GGGGCGCCCGGAGACTTGTGGGGATTCACGTGGCAAGCGCTTGGGCAAGGGGTCTGCTTGA 12192
Qy 930 ----- 930
Db 12193 GGTCCGGGTGAAGGCTGAGAGTGTGTGGGGTCCAGGTCTCGGAGAGTGTGGGGTTGGC 12252
Qy 930 ----- 930
Db 12253 TTAGGCTCAGAGATCAGAATTCACGTGAGAGATGCTCGGGGTAGGGATTATAGTTGGGG 12312
Qy 930 ----- 930
Db 12313 TCTGGGTTGGGTGCGCAGTCACTGAGGAGTGTGGGGTACTGCGCGGATTAATCTGGAATCTCT 12372
Qy 931 -----GlyMetLjuTyrLeuGlnAlaArgArgCysVal 941

```

Db 12373 CTTCAAGATCACAAGACCCCCCGAGGAGGAGTACTGCGGCTCCCGCGCTGCGG 12432  
 QY 942 HisArgaspLeuAlaIArgAsnIleLeuValGluSerGluAlaHisValysIleAla 961  
 Db 12433 CACCGCACCTGGCGCGCGGAAACATCTCTGTGAGAGGAGGACACAGCTCAAGATGCT 12492  
 QY 962 AspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGluPro 981  
 Db 12493 GACTTCGGCTAGCTAAGCTGCTGCGCTGACAAAGACTACTACTGCTGGCGCGAGCA 12552  
 QY 982 GlyIleSerProIlePheTyrTyrAlaProGluSerLeuSerAspAsnIlePheSerArg 1001  
 Db 12553 GGGCAGACCCCATTTCTTGTGTATGAGCCCGAAATCCCTCTGAGACAAATCTTCTCGG 12612  
 QY 1002 GlnSerAspValTyrPheSerPheGlyValValLeuTyrGluLeuPheTyrTyrCysAspLys 1021  
 Db 12613 CAGTCAGACGCTGTGAGCTTGGGGGCTGCTCTGTACAGAGCTCTTCACTACTGAGACAA 12672  
 QY 1022 SerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyProPro 1041  
 Db 12673 AGCTCAGACCCCTGCGCGGAGTCTCTGCGATGATGGATGATGGATGATGCTCCCGCC 12732  
 QY 1042 LeuCysArgLeuLeuGluLeuLeuAlaGluLysArgArgLeuProProProProThiCys 1061  
 Db 12733 CTCTCGCGCTCTGTGAACTGCTGAGAGGAGGCGCAGAGGCTGCGCGCGCTCTCTGCTG 12792  
 QY 1062 ProThiGluValGluGluMetGlnLeuCysTyrAlaProGluProHisAspArgPro 1081  
 Db 12793 CCGTGTGAGAGCTGACAGCTCATGAAGCTGTGCTGCGCCCTAGCCGACAGACGCGGCA 12852  
 QY 1082 AlaPheAlaThrLeuSerProGlnLeuAspProLeuTyrPheArgGlyArgProGly 1099  
 Db 12853 TCATTACGCGCTGTGGGCGCCGACGCTGACATGTGTGAGCGGAAAGCGGCGG 12906  
 RESULT 14  
 AB006011 4812 bp mRNA linear MAM 29-JUL-1997  
 LOCUS AB006011  
 DEFINITION Sus scrofa mRNA for JAK2, complete cds.  
 ACCESSION AB006011  
 VERSION AB006011.1 GI:2285795  
 KEYWORDS JAK2.  
 SOURCE Sus scrofa adult muscle thoracic aortic cell cDNA to mRNA.  
 ORGANISM Sus scrofa  
 Chordata; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Euteleostomi; Cetartiodactyla; Suidae; Sus.  
 Mammalia; Ito, Y., Mikawa, S., Kobayashi, E., Wada, Y. and Minezawa, M.  
 1 (sites)  
 Domestic pig mRNA for JAK2, complete cds  
 JOURNAL Unpublished  
 2 (bases 1 to 4812)  
 Ito, Y.  
 Direct Submission  
 Submitted (24-JUL-1997) Yoshiyasu Ito, National Institute of Animal  
 Health, Animal Genome Research Team; 2 Ikenodai, Kikuzaki-machi,  
 Inashiki-gun, Ibaragi 305, Japan (E-mail: yoshiito@nial.affrc.go.jp,  
 Tel:0298-38-8627)  
 FEATURES  
 Source  
 1. 4812  
 location/Qualifiers  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /cell\_type="thoracic aortic cell"  
 /tissue\_type="muscle"  
 /dev\_stage="adult"  
 317..3712  
 /codon\_start=1  
 /product="JAK2"  
 /protein\_id="BA21662.1"  
 /db\_xref="GI:2285796"  
 /translation="MGAACTMTMEETSTSPVQNDIGDGNANSVKIDPVLVLY  
 HSLGAEEDYLKFPAGEVVAEIVCAASKAGITPVYHSFALMNEPRLWPNHVE  
 HYDESTRNVLYRIREFPYWYCVNGSNRYHGLISRAEALVDLDFWSTLFLPROMRD  
 FLYGWVKIPIVHEOEELGMAVLDMMHIAEKQDPTLIVSSVSYFTFLPKVRAKI  
 ODYHILTKRIIRYRRIIEFQSHCKATARKLKLYLINETLQSAFYTDQFEVKEPG

RQPSGEIRATIIITNGNGIOMSRGRKHESEFTEDOLYCDPEPDIDVISIKOAND  
 GSNESRIVTIIHODKPSILKELREALSEFSLIDYUURLADAHNHLKEVAPMV  
 LENIQSNCHPILSMDFRISKILKAGNOCGLFVLRCSPDENKYLFLTEVEENYTEK  
 CLITITNENGEVLSGTNRNFSMLKDLKLCYMETVRSDSIIIFQTKCCPKPKRKN  
 LIVERINGISIDVPTSLQIRHNNVQWNRKIRNEDYIJNESLGGTFTKIKFQVRE  
 VGDYGOHETSEVLKVIDKARNYSEFPEKASMSQSLHKLIVLNCVOCGKEINTL  
 VOEFVKGSLDTYLRKKNKNSINILMLKVAOKIADAMHLEKTLIHGVACGKNITL  
 REEDRKTGNPFIKSDPGISTVLEFKDIOERILPWPEECLENRNLNATDKMSFG  
 TTLMEICSGDKPLNLDISQRLQFEDRHLQAPAKWLELANINCMIDYEDRPSF  
 RAIIRDLNLSLPTPDYELTENMDLPMRIRGALGFSGAFEDRPTQFERNLKFLQOLG  
 KNGFSYEMKRIPLDNTGCEYVAKIQHSFENLRDEFEILEILKSLQNDNIVKRY  
 GVCSASGRNRLILMEYLPYGSIRDYOLKHEERIDHKILQYTSQICKMEYIGTRKY  
 IHRDLATRNILVEMENRVLGDGFLTKLVDQKEYKKVKEPGEPSPIFWAPSLKESK  
 FSVASDVMSFGVVLYELFTYTEKSKSPAFEMFMDKQOGMYVHLEILLKNNGRIL  
 PRDPCDEIYIIMTECMNNVNOBPSFRDLALRVQDIRDMA"  
 polyA-site 4812  
 /note="17 a nucleotides"  
 BASE COUNT 1554 a 853 c 1016 g 1389 t  
 ORIGIN

Alignment scores:  
 Pzed. No.: 1..1e-152 Length: 4812  
 Score: 2637.50 Matches: 535  
 Percent Similarity: 66.15% Conservative: 194  
 Best Local Similarity: 48.55% Mismatches: 343  
 Query Match: 45.01% Indels: 30  
 DB: 4 Gaps: 14

US-09-397-967-16 (1-1099) x AB006011 (1-4812)

QY 16 CysSerLeuSerSerSerGluAlaGluAlaLeuHisValLeuLeuPro-ProArgGlyPr 35  
 Db 409 TGTGAACCAATAGATCACTCACTCTACAGCTATCTTTTCACTCCCTGGAGAACTGA 468  
 QY 35 ogLyrProGluArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysVa 55  
 Db 469 GGGG-----GATTATTGAAGTTTCCAGCTGAGAGATGTGTGCAAGAAATTTGTCT 522  
 QY 55 LARGALALALYALACYSGLYLEUProVALTYRHisSerLeuPheAlaLeuAlaTh 75  
 Db 523 TGTGCTTTTAAAGCTTGTGGAATCACCTGTGATACATAGATTTGCTTAAATGAA 582  
 QY 75 rGluAspPheSerCysTyrPheProSerHisIlePheCysIleGluAspValAspRh 95  
 Db 583 TGAAGAAGAAAGATCTGTATCCACCAACCAATGTCTTCCATGATGATGATCTACAG 642  
 QY 95 rGlnValLeuValTyrArgLeuArgPheTyrPheProAspTyrPhe-----GlyLeuGl 113  
 Db 643 GCACAATGTACTTACACAAATACATATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 702  
 QY 113 uThrCysHisArgPheGlyLeuArgGlyAspLeuThSerAlaIleLeuAspLeuHisVa 133  
 Db 703 CAGAACCTATCGCGCATGGAATATCTCGAGGTCTGAAAGCTCTTCTTCTTCTTCTTCT 762  
 QY 133 lleuGluHisIleuPheAlaGlnHisArgSerAspLeuValSerIlyArgLeuProValGl 153  
 Db 763 CARGCTTACCTTTTGTGCTGAGTGGCGACATGATTTTGTATGAGAGGTAAGATACC 822  
 QY 153 yleuSerMetLysGluGluGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMe 173  
 Db 823 TGTGATCTATGAACTCAGAGAAATGTCTTGGATGCGATGTAGATGAGAGAT 882  
 QY 173 lAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCy 193  
 Db 883 AGCCAAAGAAAGATCAAAACCCCTGACATCTATAGTTCTGTGCACCTACAGACATT 942  
 QY 193 sleuProProSerLeuArgAspValIleGlnGlyGlnAspPheValThrArgArgArg 213  
 Db 943 CTTACCAAAATGTGTTCCAGCAAAAGATCCAGACATATCATATTGTGACAGAAAGCAT 1002  
 QY 213 eArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAl 233  
 Db 1003 AAGGTACAGATTTCGACGATTATTTGAGCAATTCAGCCATTGCAAGCCACTGCCAGAA 1062

Oy 233 aleuMetalaIyTyrlleuAspleuCluArgleuHsProAlaIaIeThrThgluH 253  
 Db 1063 CTTGAACCTTAATATCTTAATCTGGAACCTTCTGCTCTCTTCTTACACAGACA 1122  
 Oy 253 rPheArGValIGlyLeuProGly-----AlaGIngluGluProGlyLeuLeuAr 269  
 Db 1123 GTTGAAGTAAAGAAAGACGGAAGAGGTCCTCAGGTGAGAGATTTTTCGAACCATTTAT 1182  
 Oy 269 gValaIaGlyAspArgnGlyIleProTrpSerSer----- 280  
 Db 1183 AATAACTGGAAACGGTGAATTCAGTGTCTCAAGAGGAAACATAAGAAAGTGAGACACT 1242  
 Oy 281 -AsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAs 300  
 Db 1243 GACAGAAACGATTTACATTTATTTGTGATTTTCTGTATTCATTTGATGTCAGTATTA 1302  
 Oy 300 nGlnAlaProArgValIGlyProAlaGlyIleuHsArgleuValThrValThrArgMetAs 320  
 Db 1303 ACAGGCAACCAAGAGGCG---TCAATATGAAGACAGCATTTGATACATTAACAGACAGA 1359  
 Oy 320 pGlyHsIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLe 340  
 Db 1360 TGGCAAAAGCTGGAAATTTGAACCTTAGCTCATTAAGAGAGAGCTTCTTTGTATCGTT 1419  
 Oy 340 uValaIaSPGlyTyrrPheArgleuIleCysAspSerArgHsIleTyrrPheCysLeuGluValAl 360  
 Db 1420 AATTGATGGATATTAGATTAATCTCAGTACATTCATTCATTTGATTAAGACATAGC 1479  
 Oy 360 aProProArgLeuLeuGluGluGluAlaAspValCysHsIleGlyProIleThrLeuAspPh 380  
 Db 1480 ACCTCCAACTGGCTTGAATATATACAAAGCAACCTGTCAAGCCCAATTTCAATGATTTT 1539  
 Oy 380 eAlaIleHsLeuLeuAlaIaIaGlySerLeuProGlyThrTyrlleuHsArgle 400  
 Db 1540 TGGCAATCAGTAACGAAAGACAGATATACAGACTGACCTTGTGTACTGATGAGCAG 1599  
 Oy 400 rProGlnAlaProArgValIGlyProGlnAlaProArgProGlyLeuProGlyLeuProAs 420  
 Db 1600 TCTTAAGACCTTAATTAATATTTTCTGACTTTTGGTGTGAGCGAAGAAATGTCACCTGA 1659  
 Oy 420 pTyrrGlyGlyCysLeuIleArgGlnAspProSerGlyAlaIaIaPheSerLeuValIGlyLeuSe 440  
 Db 1660 ATTTAAGCACTGTTTATTCACAAATAATGCAATGCGAATACACCTGAGTGTACTAG 1719  
 Oy 440 rGlnProHsArgSerLeuArgGluLeuLeuAlaIaCysTrpAsnSerGlyLeuArgVa 460  
 Db 1720 GAAGAACTTCACTAACCCTTAAGATCTTTGAAATGTTTACCAATGGAACGCTGTGCTTC 1779  
 Oy 460 IaSPGlyAlaAlaIleuTyrrLeuThrSerCysAlaProArgProGlyGluLeuSerAs 480  
 Db 1780 AGACAGATTAATTTTCCAGTTTACTTAATGCTGTCCCCCAAGCAAAAGATTAAGTCAAA 1839  
 Oy 480 nLeuIleValValArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSe 499  
 Db 1840 CCTTTAGTCTTGAACCAATGATTTCTGATGTGCCAACCTTCACCAACATTACAGAG 1899  
 Oy 499 rCysCysAlaLeuThrClnLeuSerPheHsIleProThrAspSerLeuGluTrpH 519  
 Db 1900 GCATTAATATGTCACCAATGCTGTTTCCAAATATCAGAAATGAGATTTGATATTAA 1959  
 Oy 519 sGluAsnLeuGlyHsIleGlySerPheThrTyrllePheArgGlySerIleArgGluValVa 539  
 Db 1960 TGAAGCTTGGCCAGGCAAGCGTTTACAAAATTTTAAAGGTGTAGAGAGAGAGTAGG 2019  
 Oy 539 IaSP---GlyGluThrHsAspSerGluValLeuLeuValaIaMetAspSerArgHsIle 558  
 Db 2020 AGACTAGGTCACACTGATGAACAGAGATCTTTTAAAGTTTGGATTAAGACACATAG 2079  
 Oy 558 gAsnCysMetGluSerPheLeuGluAlaIaIaSerLeuMetSerGlnValSerTyrrProH 578  
 Db 2080 AAACCTACTAGAGCTTTCTTTTGAAGCAGCAAGCATGATGAGCGCTTCTCTATTAACA 2139

Oy 578 sLeuValLeuLeuHsIleGlyValaCysMetAlaGly---AspSerIleMetValGlnGluPh 597  
 Db 2140 TCTGGTTTAAATTTATGAGATATGTGTCTGTGAGAGAGAGAAATATTCTGTTCAGAGAGT 2199  
 Oy 597 eValTyrrLeuGlyAlaIaIaLeuAspMetCysLeuArgGlyAsnIleValSerAlaSe 617  
 Db 2200 TGTAAATTTGGTCACGTGATACATCTGAAAAAGAAATTAATAATTAATATTT 2259  
 Oy 617 rTrpTyrrLeuGlnValThrTyrrGlnLeuAlaIaIaLeuAsnTyrrLeuGluAspLeuG 637  
 Db 2260 ATGGAACCTTGAAGGTGGTAAACACTTGGATGGCGCATGCAATTTTCTAGAAGACAGAAAC 2319  
 Oy 637 yLeuProHsIleGlyAsnValSerAlaArgGlyValLeuLeuAlaArgGluGly----- 655  
 Db 2320 CCTTATTCATGGGATGTGTGGCCAAAATTTCTTCTTATACAGAGAAAGACAGAGA 2379  
 Oy 655 -AspGlyAsnProProPheIleLeuSerAspProGlyValSerProThrValLeuSe 675  
 Db 2380 GACAGAAATCTCCCTTCATCAAACTTAGTATCTGTCATCAGATTAACGATTTTGCC 2439  
 Oy 675 rLeuGlnMetLeuThrAspArgIleProTrpValaIaProGlyCysLeuGlnGluAlaG 695  
 Db 2440 CAAGGACATTTGCAAGAGAGAAATACATGGGTGCTCTGTAATGATTAAGAAATCTTA 2499  
 Oy 695 nThrLeuCysLeuGluAlaAspLeuTrpGlyPheGlyAlaThrThrTrpGluValaPheG 715  
 Db 2500 AATCTAAACTTGGCAACAGATTAATGAGATTTTGGTACCCTTGTGGAAATCTGCAG 2559  
 Oy 715 nArgGlyProAlaHsIleThrSerLeuGluProAlaIaIaIaIaIaIaIaIaIaIaIaIa 735  
 Db 2560 TGGAGGTGATTAAGCCCTGAATGCTTGGATTCCTCAAAAGCAATACATCTTTTATGACAA 2619  
 Oy 735 pGlnGlyIleuProAlaLeuLeuTyrrPheGlnLeuAlaGlyLeuIleThrGlnCysWe 755  
 Db 2620 TAGGCACTGACCTTCCGCAACAAAGTGACAGACATTAACAAATTTATTAATTTGAT 2679  
 Oy 755 fAlaTyrrAspProGlyArgArgProSerPheArgAlaIaIaLeuArgAspLeuAsnGlyLe 775  
 Db 2680 GGATTTGAAGACAGATTTTAGACCTTCTTTCAGAGCCATCATACGAGCTTAATAGCTT 2739  
 Oy 775 uIleThrSerAspTyrrGlnLeuLeuSerAspProThrProGlyIleProSerProArg-- 794  
 Db 2740 GTTACTCCAGATTAAGACTACTTAACAGAA--AATGATATGTTACCAAAATGTAGCAAT 2796  
 Oy 795 -AspGluLeuCysValaIaGlyAlaGlnLeuTyrrAlaCysGlnAspProAlaIlePheG 814  
 Db 2797 AGGTGCCCTAGGCTTTCTGTGTCGCA-----TTTGAAGACCGAGACCTTAACAGTTTGA 2850  
 Oy 814 uGluArgHsIleuLeuTyrrIleSerLeuLeuGlyGlyAsnPheGlySerValIGluLe 834  
 Db 2851 AGAGACACACTTGAATAATTTCTACAGCAACTGGGCAAGGTAATTTTGGGAGTGTGAGAT 2910  
 Oy 834 uCysArgTyrrAspProLeuGlyAspAsnThrGlyProLeuValaIaValaIaIaIaIaIa 854  
 Db 2911 GTCCCGGTATGACCTCTCAAGACAAACACTGGGGAGGTGGGTGGTGAAGAAAGCTTCCA 2970  
 Oy 854 nHsSerValProAspGlnIleArgAspPheGlnArgGluIleGlnIleLeuValaIaLe 874  
 Db 2971 GCACAGTACGAGAGACACTCAGAGCTTGAAGAGAAATTTGAATCTTAATCTT 3030  
 Oy 874 uHsSerAspPheIleValaIa 894  
 Db 3031 GCAGCATATCAACATTTGTAAGTAAGAGAGGTGTCTCAAGTGTGGTGGGTATCTP 3090  
 Oy 894 uArgLeuValaIa 914  
 Db 3091 AAGATTAATTAATGAATATTACATATGGAAGTATTAACGATTAATCTTAAGAAACCTTA 3150  
 Oy 914 g---GlyLeuHsIleThrAspArgleuLeuPheAlaIaIaIaIaIaIaIaIaIaIaIaIa 933  
 Db 3151 AGAAGGATGATGATTAAGAAACTTCTTCAAGTATACATCTCAGATATGCAAGGTTATGGA 3210  
 Oy 933 uTyrrLeuGlyAlaIa 953

[illegible]



Oy	182	glyleu	leu	leu	sthr	val	ser	tyr	gly	ala	cys	leu	pro	pro	ser	leu	arg	asp	val	201		
		.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....			
Db	748	ctgcctc	ctt	gaa	ctt	gaa	ctt	gaa	ctt	gaa	ctt	gaa	ctt	gaa	ctt	gaa	ctt	gaa	ctt	gaa	907	
Oy	202	ile	gln	gln	gln	asn	phe	val	thr	arg	arg	gln	ile	arg	gln	thr	val	thr	val	leu	221	
Db	808	atccaga	gca	tat	gaa	ctt	aat	ttt	aac	tgc	gaa	agg	gat	aac	gta	ctt	ccg	agg	ctt	aat	867	
Oy	222	leu	pro	cys	ser	leu	arg	leu	pro	gly	arg	pro	tyr	ala	leu	met	ala	tyr	tyr	leu	241	
		.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....			
Db	868	cagcagt	ttc	agc	atg	gca	gaa	gca	gca	ctc	gca	gaa	gca	ctt	aaa	ctt	gaa	gtt	atc	gaa	927	
Oy	242	leu	gln	arg	leu	asn	thr	ala	thr	thr	gln	gln	thr	thr	arg	val	-----	-----	-----	256		
																	.....	.....	.....			
Db	928	ctg	gaa	gac	cc	ctc	gac	ctc	ctc	ctt	ttc	ac	gaa	gac	ggt	ttg	gaa	gaa	aat	ctc	gaa	987
Oy	257	gly	leu	pro	gly	ala	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln		
Db	988	ggg	---	cc	ttt	gac	gac	gac	gac	gac	gac	gac	gac	gac	gac	gac	gac	gac	gac	gac		
Oy	277	pro	tyr	ser	ser	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----			
Db	1045	cactg	ctc	caa	gac	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	1104	
Oy	288	phe	cys	arg	phe	pro	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln		
Db	1105	tattg	gaa	ttt	ctc	gaa	ttt	ctc	gaa	ttt	ctc	gaa	ttt	ctc	gaa	ttt	ctc	gaa	ttt	ctc	gaa	1166
Oy	308	ala	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln		
Db	1162	tca	act	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	1222	
Oy	328	phe	pro	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln	gln		
Db	1222	ctt	gac	ctc	ctt	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	gaa	1281	
Oy	348	ile	cys	arg	ser	arg	his	tyr	phe	cys	lys</											

QY	525	cyserpethlysllephedrglyserdgrdglvalvalasp---glyluthr	543
Db	1816	ggcatttttaccAAAAATATTTTAAAGGTGTAAAGAGAAAGTGGAGATTATGTCCACSTG	187
QY	544	hlsasperglvalleuleulysvalmetcassperaserghnlstargdnscymegclser	563
Db	1876	cacgaAACCGGAAGTCTTTTGAAGTCCTAGTAAAGCAACATAGAACTATTACAGACTCT	193
QY	564	phleucllnalalaserleuetserglnvalsertryrprohlsleuvalleuleuhls	583
Db	1936	ttcttttaacacacacacacATGATGATCGACTCTTCTCCAAAGCACTTGGTATTGAATTAT	199
QY	584	glyvalcysmetalaagly---aspserlmetvalnglnuphevaltrygleuiglyale	602
Db	1996	ggactatgtctctgttgacagcagcagAACATTTTGGTTCAGAGGTTTGTAAATTTGGATCC	205
QY	603	llespmetrytleuargylsargllynlstleuvalseralasertplysleuqlnval	622
Db	2056	ctggatATCAATACCTGGAAAAAATAATTCATATTAATTAATTAATTAATTAATTCAGAGT	211
QY	623	thrysglnleuvaltrylaleuansntryleuqlnsplysglyleuProhlsiglyasn	642
Db	2116	ggcaAACAGTTGGCATGGGCCATTCGACTCTCCGAGAAAAATCCCTTAATTCATGGGAAT	217
QY	643	valserAlaarglyvalleuuallaargluclyglyl-----aspglyasnpropo	660
Db	2176	gtgtgtgtccAAAAATATCTCTGCTTATCGAGAGACAGACAGAAACAGGAGATCCACCT	223
QY	661	phelleuysleuenseraprogllyalserprothryvalleuenserleuqlumetleuthr	680
Db	2236	ttcatCAAACTATGATGATCTCTGGGTTTTAGCTATCACTACATTTTACCAGAAAGCACTTTCTCAG	2295
QY	681	asparglserottryalalaproglucysleuqlnqlnlaaglnthryleucysleuqln	700
Db	2296	gagagagATCCATGGGTGACCCACTGTAGTGCATGTAACCAACCTAAATCTACTTTGGCA	2355
QY	701	AlaasplystrglyphedglvalaththtrytrgclvalpheslnargleuProhlanls	720
Db	2356	acagacAAATGGAGCTTTGGACCACTCTGTGGAGATGTGCGATGGAGAGACAAAGCT	2415
QY	721	llethrsleuqlnProalaelyslsleuysphetryglinspqlnglylnleuPro	740
Db	2416	ctgagtgctctgcactCTCAAGAAACCTGCAGTCTATAGATAGAACACCAAGCTCCCA	2475
QY	741	AlaaleuysltrphlglnleualaglyleuullethrlncysmetalaTryaspProgly	760
Db	2476	gcacCCCAAGTGACAGAGTGGCAACCTTAATACTTCACTGATGATTAAGCCACAGC	2535
QY	761	ArgargproserpheargAlaalleuualnargleuansnlyleuullethrsersaptryr	780
Db	2536	tttcagcctgccttcagacgctgctatccCGATCTAACAAGCCTCTTAACTCCAGATATAT	2595
QY	781	gluleuuleuseraspprothrproglylleproserProarg---aspqluleucysval	799
Db	2596	gaacTcGTGCAGAGAA---AATGACATGCTCCCAACATGAGAAATAGTCCCTTAGATTT	2652
QY	800	Alaaglyalaglnleuylalacysglnaspproalaallephelnglnlarnghlsleuys	819
Db	2653	ttctgtctgt-----tttGAAGACAGGGAGCCCTACACATTTTGAAGAGACACCTTGA	2706
QY	820	tyrllserleuileuqllysglyasnphelgyservAlaGluencysargtryrAspro	839
Db	2707	ttttctacagCAACTTGGCAAGGGTAACTTTGGGAGTGTGAGATGTGCCCTATGACCCG	2766
QY	840	leuqlnyspansntrglyProleuvalalavalysglnleuqlnhlserservalProasp	859
Db	2767	cttcagagcACAACCTGGAGGAGGTGATGCTTTAAAGAGCTCCAGACAGCAGCGAAGAG	2826
QY	860	glnqlnargspheqlnargqlnleqlnleuylsAlaaleuylhlserserAspheille	879
Db	2837	catctccgacagcattgAAAGGAAATTTGAATCTCGAAATCTTGGACATGACAACT	2886
QY	880	Vallystrylarginglylsertryrcllyrproglylarginleuvalleuvalmetgln	899



GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 28, 2003, 13:46:19 ; Search time 105.732 Seconds  
(without alignments)  
4679.258 Million cell updates/sec

Title: US-09-397-967-15  
Perfect score: 90  
Sequence: 1 AKLPLDKDYVVRPEP 17

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_p2n.model -DEV=xl  
-Q=/cgn2\_1/USPTO.spool/US09397967/rnat\_26042003.182312.10714/app.query.fasta.1.1486  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPTCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human4.cdt -LIST=45  
-DOCFALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pic -MORF=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967.ecgn.1.1.5398.etrnat\_26042003.182312.10714 -NCPU=6 -ICPU=3  
-NO\_XLPRY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
GenEmbl :  
1: gb\_ba : \*  
2: gb\_hlg : \*  
3: gb\_in : \*  
4: gb\_ov : \*  
5: gb\_ov : \*  
6: gb\_pat : \*  
7: gb\_ph : \*  
8: gb\_pl : \*  
9: gb\_pr : \*  
10: gb\_ro : \*  
11: gb\_srs : \*  
12: gb\_sy : \*  
13: gb\_un : \*  
14: gb\_vl : \*  
15: em\_ba : \*  
16: em\_fun : \*  
17: em\_hum : \*  
18: em\_in : \*  
19: em\_mu : \*  
20: em\_om : \*  
21: em\_or : \*  
22: em\_ov : \*  
23: em\_ph : \*  
24: em\_pl : \*  
25: em\_ro : \*  
26: em\_srs : \*  
27: em\_un : \*  
28: em\_un : \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	165	6 E08798	E08798 cDNA encodl
2	90	100.0	198	6 HSU08340	U08340 Human clone
3	90	100.0	448	6 AX203080	AX203080 Sequence
4	90	100.0	778	6 AX203087	AX203087 Sequence
5	90	100.0	2914	6 AX203084	AX203084 Sequence
6	90	100.0	3620	6 HSU031601	U31601 Human tyros
7	90	100.0	4064	6 AX203081	AX203081 Sequence
8	90	100.0	4064	6 HSU09607	U09607 Human JAK f
9	90	100.0	13562	9 HSU070065	U70065 Human JAK3
10	90	100.0	21323	9 AFS13860	AFS13860 Homo sapi
11	90	100.0	41006	9 AC007201	AC007201 Homo sapi
12	83	92.2	3528	10 M05JAK3A	L40172 Mus musculu
13	83	92.2	3723	10 M05PYK3N	L32955 Mouse prote
14	83	92.2	4016	10 M05JAK3H	L33768 Mus musculu
15	83	92.2	10326	10 MMU071201	U71201 Mus musculu
16	83	92.2	218208	2 AC073700	AC073700 Mus muscu
17	83	92.2	223734	2 AC073750	AC073750 Mus muscu
18	80	88.9	3644	5 AF034576	AF034576 Gallus ga
19	76	84.4	3778	10 RAT04AK3	D28508 Rat mRNA fo
20	71	78.9	14907	5 AF090382	AF090382 Tetradon
21	69	76.7	4387	5 DRES690	AJ005690 Danio rer
22	66	73.3	171	5 XU011724	U11724 Xenopus lae
23	64	71.1	166	5 AB062055	AB062055 Gallus ga
24	64	71.1	214	10 M05PTK	M33423 Mouse prote
25	64	71.1	1967	5 DRES691	AJ005691 Danio rer
26	64	71.1	3435	6 AR129823	AR129823 Sequence
27	64	71.1	3495	6 AR048200	AR048200 Sequence
28	64	71.1	3495	6 AR068116	AR068116 Sequence
29	64	71.1	3495	6 AR071176	AR071176 Sequence
30	64	71.1	3495	6 AR177639	AR177639 Sequence
31	64	71.1	3495	6 162349	162349 Sequence 2
32	64	71.1	3495	6 187981	187981 Sequence 2
33	64	71.1	3500	6 AF001362	AF001362 Homo sapi
34	64	71.1	3629	6 AR136148	AR136148 Sequence
35	64	71.1	3629	6 AR143945	AR143945 Sequence
36	64	71.1	3629	6 192575	192575 Sequence 8
37	64	71.1	3629	10 M05PTK	L13956 Mouse prote
38	64	71.1	3731	10 RN013396	AB036337 Sus scro
39	64	71.1	4042	9 AB036337	AB036337 Homo sapi
40	64	71.1	4161	9 AF005216	AB036336 Sus scro
41	64	71.1	4599	4 AB036336	AB036336 Sus scro
42	64	71.1	4812	4 AB006011	AB006011 Sus scro
43	64	71.1	5117	6 AX354568	AX354568 Sequence
44	64	71.1	5117	6 AF058925	AF058925 Homo sapi
45	64	71.1	85671	2 AC096324_3	Continuation (4 of

ALIGNMENTS

RESULT 1

E08798 165 bp RNA linear PAT 29-SEP-1997  
 LOCUS E08798 cDNA encoding novel human tyrosine kinase.  
 DEFINITION E08798  
 ACCESSION E08798.1 GI:2176910  
 VERSION JP 1995059569-A/1.  
 KEYWORDS Homo sapiens.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 165)  
 AUTHORS Sakano, S.  
 TITLE NOVEL TYROSINE KINASE PEPTIDE AND DNA CODING THE SAME  
 JOURNAL Patent: JP 1995059569-A 1 07-MAR-1995  
 ASAHIT-CHEM-IND CO LTD  
 OS Homo sapiens (human)  
 PN JP 1995059569-A/1  
 PD 07-MAR-1995  
 PFI 25-AUG-1993 JP 1993210404  
 K12 SAKANO SEIJI  
 PC C12N15/09,C12N9/12;  
 CC strandedness: Double;  
 CC topology: Linear;  
 CC hypothetical: No;  
 CC anti-sense: No;  
 FH Key Location/Qualifiers  
 FT source 1.165  
 FT /organism="Homo sapiens"  
 FT /cell\_type="megakaryoblast"  
 FT /cell\_line="UT-7"  
 FT 1.165  
 FT mat.peptide /product="novel human tyrosine kinase".  
 FEATURES  
 source Location/Qualifiers  
 1.165  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 36 a 56 c 39 g 34 t  
 ORIGIN  
 Alignment Scores: 9.04e-09 Length: 165  
 Pred. No.: 90.00 Matches: 17  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 100.00% Gaps: 0  
 DB: 6  
 US-09-397-967-15 (1-17) x E08798 (1-165)  
 Oy 1 AlalysleuLeuPProleuAspLysAspTyrTyrValValArgGluProGly 17  
 Db 52 GCTAAGCTGCTGCCGCTTGACAAAGACTACTAGCTGTCGCGACGACGCC 102  
 RESULT 2  
 LOCUS HSU08340 198 bp mRNA linear  
 DEFINITION Human clone NTK16 tyrosine kinase mRNA, partial cds.  
 ACCESSION U08340  
 VERSION U08340.1 GI:473879  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 198)  
 AUTHORS Fuortes, M.  
 TITLE Tumor Necrosis Factor and Adhesion: Joint Control of Neutrophil  
 JOURNAL Activation  
 Thesis (1994) Cell Biology and Anatomy, Cornell University,  
 Graduate School of Medical Sciences  
 REFERENCE 2 (bases 1 to 198)  
 AUTHORS Fuortes, M.

Yes

PAT 21-APR-1994

TITLE Direct Submission  
 JOURNAL Submitted (05-APR-1994) Michele Fuortes, Cell Biology and  
 Anatomy/Medicine, Cornell University Medical College, 1300 York  
 Avenue, New York, NY 10021, USA  
 FEATURES  
 source Location/Qualifiers  
 1.198  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="NTK16"  
 /cell\_type="neutrophil"  
 /tissue\_type="blood"  
 <1..>198  
 /codon\_start=1  
 /evidence=not\_experimental  
 /product="tyrosine kinase"  
 /protein\_id="AAAI1743.1"  
 /db\_xref="GI:473880"  
 /translation="DLARNILVESAHKINDPGLAKLLPDKDYVVERGQSPIF  
 WYAFSLSDNIFRSQSDVWVSG"  
 CDS  
 BASE COUNT 41 a 66 c 52 g 39 t  
 ORIGIN  
 Alignment Scores: 1.1e-08 Length: 198  
 Pred. No.: 90.00 Matches: 17  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 100.00% Gaps: 0  
 DB: 9  
 US-09-397-967-15 (1-17) x HSU08340 (1-198)  
 Oy 1 AlalysleuLeuPProleuAspLysAspTyrTyrValValArgGluProGly 17  
 Db 67 GCTAAGCTGCTGCCGCTTGACAAAGACTACTAGCTGTCGCGACGACGCC 117  
 RESULT 3  
 LOCUS AX203080 448 bp DNA linear PAT 30-AUG-2001  
 DEFINITION Sequence 1 from Patent W00152892.  
 ACCESSION AX203080  
 VERSION AX203080.1 GI:15392437  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 448)  
 AUTHORS Vasilos, G.  
 TITLE Jak/stat pathway inhibitors and the uses thereof  
 JOURNAL GENZYME CORPORATION (US)  
 FEATURES  
 source Location/Qualifiers  
 1.448  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 79 a 150 c 127 g 92 t  
 ORIGIN  
 Alignment Scores: 2.63e-08 Length: 448  
 Pred. No.: 90.00 Matches: 17  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 100.00% Gaps: 0  
 DB: 6  
 US-09-397-967-15 (1-17) x AX203080 (1-448)  
 Oy 1 AlalysleuLeuPProleuAspLysAspTyrTyrValValArgGluProGly 17  
 Db 306 GCTAAGCTGCTGCCGCTTGACAAAGACTACTAGCTGTCGCGACGACGCC 356

RESULT 4  
AX203087 778 bp DNA linear PAT 30-AUG-2001  
LOCUS Sequence 8 from Patent WO0152892.  
DEFINITION AX203087  
ACCESSION AX203087  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Vassios, G.  
TITLE Jak/stat pathway inhibitors and the uses thereof  
JOURNAL Patent: WO 0152892-A 8 26-JUL-2001;  
GENZYME CORPORATION (US)  
FEATURES  
source  
1. 778  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 131 a 265 c 219 g 163 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.75e-08 Length: 778  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-397-967-15 (1-17) x AX203087 (1-778)  
Oy 1 AlalysleuleuProleuasplysAsptyTyTyValValargLupProgly 17  
|||||  
Db 187 GCTACCTGCTGCCGCTTGACAAAGACTACTACTGCTGCCGACGACGC 237  
RESULT 5  
AX203084 2914 bp DNA linear PAT 30-AUG-2001  
LOCUS Sequence 5 from Patent WO0152892.  
DEFINITION AX203084  
ACCESSION AX203084  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Vassios, G.  
TITLE Jak/stat pathway inhibitors and the uses thereof  
JOURNAL Patent: WO 0152892-A 5 26-JUL-2001;  
GENZYME CORPORATION (US)  
FEATURES  
source  
1. 2914  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 566 a 907 c 809 g 632 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.95e-07 Length: 2914  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-397-967-15 (1-17) x AX203084 (1-2914)  
Oy 1 AlalysleuleuProleuasplysAsptyTyTyValValargLupProgly 17  
|||||  
Db 1856 GCTACCTGCTGCCGCTTGACAAAGACTACTACTGCTGCCGACGACGC 1906

RESULT 6  
HSU31601 3620 bp mRNA linear PR1 24-OCT-1995  
LOCUS Human tyrosine protein kinase (JAK3b) splice variant mRNA, complete  
DEFINITION cds.  
U31601  
ACCESSION U31601.1 GI:1039418  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Lai, K.S., Jin, Y., Graham, D.K., Witthuhn, B.A., Ihle, J.N. and  
Liu, E.T.  
TITLE A kinase-deficient splice variant of the human JAK3 is expressed in  
hematopoietic and epithelial cancer cells  
JOURNAL J. Biol. Chem. 270 (42), 25028-25036 (1995)  
MEDLINE 96027605  
PUBMED 7559633  
REFERENCE  
AUTHORS Witthuhn, B.A., Silvennoinen, O., Mura, O., Lai, K.S., Cwik, C.,  
Liu, E.T., and Ihle, J.N.  
TITLE Involvement of the Jak-3 Janus kinase in signalling by interleukins  
2 and 4 in lymphoid and myeloid cells  
JOURNAL Nature 370 (6485), 153-157 (1994)  
MEDLINE 8022486  
PUBMED  
REFERENCE  
AUTHORS 3 (bases 1 to 3620)  
Liu, E.T.  
TITLE Direct Submission  
JOURNAL Submitted (14-JUL-1995) Edison T. Liu, Lindeberger Comprehensive  
Cancer Center, University of North Carolina, Chapel Hill, NC  
27599-7295, USA  
Witthuhn, B.A. Nature 370, 153-157, 1994.  
FEATURES  
source  
1. 3620  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19p12"  
/clone="SKBR-3"  
/cell\_type="SKBR-3 breast cancer cells"  
1. 3620  
/gene="JAK3b"  
96. 3380  
/gene="JAK3b"  
/function="tyrosine protein kinase"  
/codon\_start=1  
/product="JAK3b"  
/protein\_id="AAC50226.1"  
/db\_xref="GI:1039419"  
/translation="MAPSPSEPLIPORSCSLSTFAGALHVLPPARGGPORLTSFS  
FGDLAEDLCVOAKASGILPYHSILFALATEDLSMPESHFSVEDASTOVLARI  
RRTFPMNFGLEKRRGRLKRDLSALDLPVLEHLPFAQRHSDLSYRRLRRTVR  
ECUSLAVLDLARARQAOORPGLTAKLTVSKALPSPRLDIOGLSFVTRRRIKTVR  
RALRRVAAQADNRHSLMAKYMIDLERLDPAEAGETRVHGLGSHDFVGLRVAGD  
GGIAWQEQEQLVDFPCDEPEIVDISIKQAPRVGAPAGELRVLTFTDNOILAEKRG  
LPEALSFVALVDGFRLLTSDHFECEIVAPRLLEVEAEOGCHPTLDFAIKRLTG  
GSRPQSVILRRSPDDPSFLITVCVONPGLDGKGLIRSPGTGTELLVSPHSLSL  
RELATLACWDSGLHVDGAVYLTSCCIPRPREKSNLIVONGHSPPTSSLVPOSOYOL  
SOMTFKRTIPADSLMEHNLGHSFTKTYRCGRHEVDGARKTEVLAKWDAAKHKCM  
ESLELASLSQVSRHLVYLHGVCMAGDSLVQEFVHLAIDYLRKRGHLVPASMK  
LOYVVKLALVNLLEDKGLPHGVNSARKVLLAREGADGSPPTKLDPGSPAVLSIE  
MLDRIIPWVAPFCLREAOITSLADKKMGFGATVWEVSGVTPMISALDPAKIKQFED  
ROOLPAKPMELALLIQCCMAVEPVORSPRAYIRDLNSLISDPTLSPPTGALAP  
RDLGMAQOILYACODPTIFERHLKYSQCGKNGFSGELCRDPTLGDNGALVAVKO  
LQHSQPDQORDFOREITILKALHSDFTVTKRGYSTGFGROSLVMEYTPDSCGLRPL  
QRRRARLDARLLYSQITCKMEYLGSRVCVRDLAARNLIVSEAHVA IADGLAK  
LPLDKDYVYVREPQSPITFWAPESLSNIFRSQDSVFGVALLFLFYCDKSCSP  
SAFLRMGCEBVPALCRILLELEEGORLIPADPACPAAEVSAAGLASVSOSVDMAGVS  
GKPAQA"

BASE COUNT 670 a 1152 c 1076 g 722 t  
 ALIGNMENT SCORES:  
 Pred. No.: 2,466-07 Length: 3620  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-397-967-15 (1-17) x HSU31601 (1-3620)

OY 1 AlalysleuLeuProleuAspLysAspTyrTyValValArgLupProgly 17  
 DB 3006 GCTAAGCTGCTGCCGCTTGACAAAGACTACTACGTGCTCCGCGAGCCAGC 3056

RESULT 7  
 LOCUS AX203081 4064 bp DNA linear PAT 30-AUG-2001  
 DEFINITION Sequence 2 from Patent WO0152892.  
 ACCESSION AX203081  
 VERSION AX203081.1 GI:15392438  
 KEYWORDS  
 SOURCE human  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 4064)  
 AUTHORS Vasio, G.  
 TITLE Jak/stat pathway inhibitors and the uses thereof  
 JOURNAL Patent: WO 0152892-A 2-26-JUL-2001;  
 GENZYME CORPORATION (US)  
 FEATURES  
 source Location/Qualifiers  
 1..4064  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 746 a 1292 c 1168 g 858 t  
 ALIGNMENT SCORES:  
 Pred. No.: 2,786-07 Length: 4064  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-397-967-15 (1-17) x AX203081 (1-4064)

OY 1 AlalysleuLeuProleuAspLysAspTyrTyValValArgLupProgly 17  
 DB 3006 GCTAAGCTGCTGCCGCTTGACAAAGACTACTACGTGCTCCGCGAGCCAGC 3056

RESULT 8  
 LOCUS HSU09607 4064 bp mRNA linear PRI 09-JUL-1994  
 DEFINITION Human JAK family protein tyrosine kinase (JAK3) mRNA, complete cds.  
 ACCESSION U09607  
 VERSION U09607.1 GI:508730  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 1124)  
 AUTHORS Kawamura, M.; McGivicar, D.W.; Johnston, J.A.; Blake, T.B.; Chen, Y.;  
 O'Shea, J.; Thygesen, A.B.; Kelly, D.; Stimples, J.E.; Ortaldo, D.R. and  
 Lesh, B.K.  
 TITLE Molecular cloning of L-JAK, a Janus family protein-tyrosine kinase  
 expressed in natural killer cells and activated leukocytes  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (14), 6374-6378 (1994)  
 MEDLINE 94291384

PUBMED 8022790  
 REFERENCE 2 (bases 1 to 4064)  
 O'Shea, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-MAY-1994) John O'Shea, Leukocyte Cell Biology Section  
 LBI BMD, National Cancer Institute FCRCO, Bldg 560 Rm 3146 FCRCO,  
 Frederick, MD 21702, USA  
 FEATURES  
 source Location/Qualifiers  
 1..4064  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="Yr, HUT-78"  
 /cell\_type="natural killer cell, PHA activated T cells"  
 1..4064  
 /gene="JAK3"  
 96..3470  
 /gene="JAK3"  
 /codon\_start=1  
 /product="JAK family protein tyrosine kinase"  
 /protein\_id="AA19626.1"  
 /db\_xref="GI:508731"

CDS  
 gene  
 CDS

BASE COUNT 746 a 1292 c 1168 g 858 t  
 ALIGNMENT SCORES:  
 Pred. No.: 2,786-07 Length: 4064  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-397-967-15 (1-17) x HSU09607 (1-4064)

OY 1 AlalysleuLeuProleuAspLysAspTyrTyValValArgLupProgly 17  
 DB 3006 GCTAAGCTGCTGCCGCTTGACAAAGACTACTACGTGCTCCGCGAGCCAGC 3056

RESULT 9  
 LOCUS HSU70065 13562 bp DNA linear PRI 27-JAN-1997  
 DEFINITION Human JAK3 gene, complete cds.  
 ACCESSION U70065  
 VERSION U70065.1 GI:1800224  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 13562)  
 AUTHORS Riedy, M.C.; Dutra, A.S.; Blake, T.B.; Modi, W.; Lal, B.K.; Davis, J.;  
 Bosse, A.; O'Shea, J.J. and Johnston, J.A.  
 TITLE Genomic sequence, organization, and chromosomal localization of



```

        intron          11257..12008
                        /gene="JAK3"
                        /number=17
        exon            12009..12125
                        /gene="JAK3"
                        /number=18
        intron          12126..12399
                        /gene="JAK3"
                        /number=18
        exon            12400..13562
                        /gene="JAK3"
                        /number=19

BASE COUNT      2939 a      3476 c      4009 g      3131 t      7 others
ORIGIN

Alignment Scores:
Pred. No.:      1.01e-06      Length:      13562
Score:          90.00      Matches:      17
Percent Similarity: 100.00%      Conservative: 0
Best local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              9      Gaps:      0

US-09-397-967-15 (1-17) x HSU70065 (1-13562)

OY      1      AlalysleuLeuProLeuAspLysAspTyrValValAlaArgLupProGly 17
        |||||
Db      12505 GCTAAGCTGCTGCCGCTTGACAAAGACTACTACGTGCTGCCGAGCCAGGC 12555

RESULT 10
AF513860      21323 bp      DNA      linear      PRI 09-JUL-2002
LOCUS      Homo sapiens Janus kinase 3 (a protein tyrosine kinase, leukocyte)
DEFINITION      (JAK3) gene, complete cds.
ACCESSION      AF513860
VERSION      AF513860.1      GI:21263105
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 21323)
AUTHORS      Rieder,M.-J., Armet,F.Z., Carrington,D.P., Ozuna,M., Kuldanek,S.A.,
        Rajkumar,N., Toch,E.J., Yi,Q. and Nickerson,D.A.
        Direct Submission
        Submitted (20-MAY-2002) Genome Sciences, University of Washington,
        1705 NE Pacific, Seattle, WA 98195, USA
        To cite this work please use: SeattleSNPs, NHLBI HL66682 Program
        for Genomic Applications, UW-PHRC, Seattle, WA (URL:
        http://pga.gs.washington.edu).
FEATURES
        source
            1..21323
                Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
        repeat_region
            1..214
                /rpt_family="Alu"
        variation
            290
                /rpt_type=dispersed
                /frequency="0.13"
                /replace="a"
        variation
            370
                /frequency="0.02"
                /replace="t"
        variation
            575
                /frequency="0.33"
                /replace="a"
        variation
            622
                /frequency="0.01"
                /replace="a"
        repeat_region
            673..965
                /rpt_family="Alu"
        variation
            681
                /rpt_type=dispersed

        /frequency="0.02"
        /replace="a"
        784
        /frequency="0.13"
        /replace="a"
        925
        /frequency="0.33"
        /replace="g"
        970
        /frequency="0.01"
        /replace="t"
        994
        /frequency="0.96"
        /replace="t"
        1019..1270
        /rpt_family="Alu"
        /rpt_type=dispersed
        1039
        /frequency="0.04"
        /replace="c"
        1051
        /frequency="0.01"
        /replace="a"
        1061
        /frequency="0.06"
        /replace="t"
        1064
        /frequency="0.04"
        /replace="t"
        1273..1336
        /rpt_family="Alu"
        /rpt_type=dispersed
        1337..1630
        /rpt_family="Alu"
        /rpt_type=dispersed
        1467
        /frequency="0.30"
        /replace="a"
        1513
        /frequency="0.04"
        /replace="t"
        1564
        /frequency="0.23"
        /replace="a"
        1590..2091
        /note="Region not scanned for variation"
        1631..1706
        /rpt_family="Alu"
        /rpt_type=dispersed
        1966..2274
        /rpt_family="Alu"
        /rpt_type=dispersed
        2359..20650
        /gene="JAK3"
        join(<2359..2557,2891..3014,3300..3411,3619..3764,
        join(2374..2557,2891..3014,3300..3411,3619..3764,
        4181..4475,5029..5151,5245..5402,6455..6566,7133..7319,
        8406..8533,8733..8864,9583..9667,10746..10873,
        11582..11714,11794..11945,12076..12226,13868..14007,
        14089..14278,14999..15123,15397..15569,16178..16295,
        16580..16690,19892..20050)
        /gene="JAK3"
        /product="Janus kinase 3 (a protein tyrosine kinase,
        leukocyte)"
        join(2374..2557,2891..3014,3300..3411,3619..3764,
        4181..4475,5029..5151,5245..5402,6455..6566,7133..7319,
        8406..8533,8733..8864,9583..9667,10746..10873,
        11582..11714,11794..11945,12076..12226,13868..14007,
        14089..14278,14999..15123,15397..15569,16178..16295,
        16580..16690,19892..20050)
        /gene="JAK3"
        /codon_start=1
        /product="Janus kinase 3 (a protein tyrosine kinase,
        leukocyte)"

```



```

/protein_id="AA044860.1"
/db_xref="GI:21263106"
/translation="MAPSEETPLIPORSCSLSTREAGALHYLLPARGGPPORLSFS
RQDHLAEDLCNOAKASGTLIPYHSLFALATPDLSCMPRSSISVEASTOYLXRI
RRTFPMFGLCKHFRGLKDLASILDLPIYEHLPQHRSDLVSGRLVGLSLKEOG
ECLSLAVLDLARMAREQARPEGLKTLVSKCLPSLSLDLQGLFVYRRRIKRTVR
RALRVACQADRHSLMAKYIMDLERLPAGAAEFHVLPGALGHDLGLLRVAGD
CGIANTOGCEVLOPFCDFPEVIDISIKOAPVAGGHEHLYVTRTDQGLLEAFEG
LPEALSEVALVDGFRLLTDSQHFCKEVAAPRLLEVEAGCHGPTLDLPAINKLKTG
GSRGCSYLRSSRDPDFSLTVCVONPIGPDYKGLIRSPGTGHEHLYVGRSISL
RELLATCMDCGLHNDGVATLTSCCIPRKEKSNLIVORGSPTSSLVQSOYQL
SOMTHKIPADSLHNEHNGHSGFTKITRGCHVEVVDGARKTEVLKVMKHNKM
ESFLBASLSQSVYRHLVLCVMAAGDSTVOEHLGALIDMYLRKRGHLPVAPSK
LOVYQOLAYALNYLEDKGLPHGNVSAKRYLAREGADGSPFIKLSDPGSPVLSLE
MLTDRIIPMYAPECLEAQTLSLEADKMGAGVAYEVESGVTMPSALDPARKLOPED
ROOLPAPKMTLALLIOCMAYEPVORSEFRAVITDINSLISSDVELLSDTPGALAP
RDGLMNGAOLYACDPTIFEEHKLKYSQCGKNGSVLCLKDPDGLDWTGALVAVKO
LOHSGPDQORDFOREIQLKALHSDPIYAKVSTGPGKQSLRLWEYVLPSCGLRDL
QNHRRALDASRLLYSSQICKGMEYLGSRVCRVRODLAANNILVESEHAKIADFGIAK
LPLDKDYVYVREPQSPIFWYAPESLSNIPFSRQSDPVASFVLYELFTYCDKSCSP
SAEFLRMNGCEDVPALCRLELLEEGORLPAPPCAPVHMLKLCNAPSPQDRPSE
SALGPQDLMLMNGSGRCETHAFTAHPECKHNSLFS"
variation
2579
/gene="JAK3"
/frequency="0.36"
/replacement="t"
2597
/gene="JAK3"
/frequency="0.43"
/replacement="a"
2599
/gene="JAK3"
/frequency="0.37"
/replacement="a"
2619.2778
/rpt_family="L2"
/rpt_type="dispersed"
variation
2653
/gene="JAK3"
/frequency="0.32"
/replacement="t"
2654
/gene="JAK3"
/frequency="0.04"
/replacement="a"
3003
/gene="JAK3"
/frequency="0.07"
/replacement="c"
3030
/gene="JAK3"
/frequency="0.02"
/replacement="a"
3076
/gene="JAK3"
/frequency="0.07"
/replacement="a"
3095.3200
/rpt_family="MTR"
/rpt_type="dispersed"
variation
3106
/gene="JAK3"
/frequency="0.06"
/replacement="a"
3186
/gene="JAK3"
/frequency="0.06"
/replacement="c"
3187
/gene="JAK3"
/frequency="0.06"
/replacement="a"
3211
variation

```

US-09-397-967-15 (1-17) x AF513860 (1-21323)

Alignment Scores:  
 Pred. No.: 1,64e-06 Length: 21323  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

QY 1 AlalysleuLeuProleuAsplysAspTyTyrValValArgGluProGly 17  
 Db 15502 GCTAACGCTGCTGCCGCTTGACAAAGACTACTACGTGCTGCCGACGACGCG 15552

RESULT 11  
 AC007201/c 41006 bp DNA linear PRI 03-APR-1999  
 LOCUS Homo sapiens chromosome 19, cosmid R34383, complete sequence.  
 DEFINITION AC007201  
 ACCESSION AC007201  
 VERSION AC007201.1 GI:4558769  
 KEYWORDS HTG.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Lamerdin,J.E., McCready,P.M., Skowronski,E., Scott,D.,  
 Burkhardt-Schultz,K., Gordon,L., Dias,J., Sakadatsis,G.,  
 Stillwagen,S., Phan,H., Velasco,N., Do,L., Regala,W., Terry,A.,  
 Dangnanan,L., Erler,A., Christensen,M., Georgescu,A., Avila,J.,  
 Atlix,C., Andreise,T., Amico-Keller,G., Coefield,J., Duarte,S.,  
 Lucas,S., Bruce,R., Thomas,P., Quan,G., Krommiller,B., Arellano,A.,  
 Sanders,C., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S. and  
 Carrano,A.V.  
 Sequence analysis of a 5.7 Mb region in 19p13.1

TITLE  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 41006)  
 AUTHORS Lamerdin,J.E.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-APR-1999) Joint Genome Institute, Lawrence Livermore  
 National Laboratory, 7000 East Ave., Livermore, CA 94551, USA

COMMENT  
 Map and sequence oriented from p telomere to centromere. Cosmid  
 R34383 overlaps cosmid F19847 (AC005952) to the left from bases 1  
 to 2,923 of this accession, and overlaps cosmid R31408 (AC005796)  
 to the right from bases 37,823 to 41,006. Additional chromosome 19  
 map and sequence information are available at:  
 http://www-bio.llnl.gov/dbp/genome/genome.html.

FEATURES  
 Location/Qualifiers  
 1. 41006  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /map="19p13.1 between JAK3 and PDE4C"  
 /cell\_line="5HL2-B"  
 /clone="R34383"  
 /clone\_lib="LL19NC03 R chromosome 19-specific cosmid  
 library"

/note="Cosmid library constructed at LLNL from flow-sorted  
 chromosomes from human-hamster hybrid 5HL2-B, which  
 carries chromosome 19 as its only human chromosome."

```

repeat_region      1. .67
/rpt_family="AluY/monomer"
repeat_region      143. .441
/rpt_family="AluYb"
repeat_region      complement(578. .885)
/rpt_family="AluX"
CDS                complement(<1078. .1272)
/note="LEYDIG INSULIN-LIKE PEPTIDE PRECURSOR"
/codon_start=1
/product="INL3 HUMAN [AA 1- 65]"
/protein_id="AAD22740.1"
/db_xref="GI:4558770"
/translation="MDPRLPAWALVLLGPALVAFALGAPPEMREKLCGHFVAVLR
VCGGPRMSTEARPATGCDRE"
repeat_region      complement(1486. .1786)
/rpt_family="AluX"
repeat_region      complement(1803. .1970)
/rpt_family="AluSg/x"
repeat_region      complement(1975. .2275)
/rpt_family="AluYb"
repeat_region      complement(2279. .2444)
/rpt_family="AluYb"
repeat_region      complement(2447. .2747)
/rpt_family="AluX"
repeat_region      complement(2748. .2877)
/rpt_family="AluYb"
repeat_region      3052. .3357
/rpt_family="AluX"
repeat_region      complement(3365. .3412)
/rpt_family="MIR"
repeat_region      complement(3773. .3846)
/rpt_family="(GA)n"
repeat_region      complement(3851. .4142)
/rpt_family="AluX"
repeat_region      complement(4914. .4995)
/rpt_family="MIR"
repeat_region      5386. .5686
/rpt_family="AluY"
repeat_region      5739. .5848
/rpt_family="FLAM_A"
repeat_region      5870. .6170
/rpt_family="AluX"
repeat_region      complement(5918. .24239)
/gene="JAK3"
/note="Janus kinase 3 (a protein tyrosine kinase,
leukocyte)"
mRNA               complement(join(5918. .6676,9874. .9984,10269. .10386,
10994. .11166,11440. .11564,12285. .12474,12556. .12695,
14337. .14487,14618. .14769,14849. .14981,15690. .15817,
16885. .16979,17698. .17829,18029. .18156,19243. .19429,
19986. .20107,21155. .21312,21406. .21528,22082. .22376,
22793. .22938,23146. .23257,23543. .23666,24000. .24239))
/gene="JAK3"
/product="Human JAK family protein tyrosine kinase (JAK3)
mRNA, complete cds."
repeat_region      complement(6416. .6472)
/rpt_family="MIR"
CDS                complement(join(6509. .6676,9874. .9984,10269. .10386,
10994. .11166,11440. .11564,12285. .12474,12556. .12695,
14337. .14487,14618. .14769,14849. .14981,15690. .15817,
16885. .16979,17698. .17829,18029. .18156,19243. .19429,
19986. .20107,21155. .21312,21406. .21528,22082. .22376,
22793. .22938,23146. .23257,23543. .23666,24000. .24239))
/gene="JAK3"
/note="TYROSINE-PROTEIN KINASE JAK3; JANUS KINASE 3;
JAK-3; LEUKOCYTE JANUS KINASE; L-JAK"
/codon_start=1
/product="JAK3_HUMAN"
/protein_id="AAD22741.1"
/db_xref="GI:4558771"
/translation="MAPSEETPLIPORSCSLSTEAGALHVLIPARGSPORLUSP
FGHLADLCVOAKASGILPYVHSIFALATEDISCFPPSHITSVEDASQVLLIRI
RFIFPNWFGLEKCHRFGLRKDLASALLDPLVLEHLFAOHNSDLVSGRLPVGLSLKEOG

```

```

repeat_region      complement(6747. .7040)
/rpt_family="AluX"
repeat_region      complement(7111. .7412)
/rpt_family="AluSg"
repeat_region      7441. .7688
/rpt_family="AluSp"
repeat_region      complement(7985. .8020)
/rpt_family="(CAAA)n"
repeat_region      complement(8021. .8315)
/rpt_family="AluYb"
repeat_region      complement(8316. .8454)
/rpt_family="AluYb"
repeat_region      8499. .8564
/rpt_family="LINE2"
repeat_region      8643. .8902
/rpt_family="AluYb"
repeat_region      complement(8929. .9224)
/rpt_family="AluX"
repeat_region      9243. .9536
/rpt_family="AluYb"
repeat_region      complement(9560. .9753)
/rpt_family="MIR"
repeat_region      11759. .12044
/rpt_family="AluX"
repeat_region      complement(12074. .12154)
/rpt_family="MIR"
repeat_region      13194. .13323
/rpt_family="FLAM_A"
repeat_region      13337. .13638
/rpt_family="AluX"
repeat_region      complement(13651. .13805)
/rpt_family="AluYb"
repeat_region      13832. .14140
/rpt_family="AluX"
repeat_region      complement(16047. .16337)
/rpt_family="AluX"
repeat_region      complement(17180. .17480)
/rpt_family="AluX"
repeat_region      complement(17487. .17609)
/rpt_family="FLAM_C"
repeat_region      complement(17910. .17976)
/rpt_family="MIR"
repeat_region      complement(18301. .18479)
/rpt_family="AluSg/x"
repeat_region      complement(18480. .18780)
/rpt_family="AluX"
repeat_region      complement(19505. .19804)
/rpt_family="AluSc"
repeat_region      20248. .20548
/rpt_family="AluY"
repeat_region      complement(20590. .20882)
/rpt_family="AluYb"
repeat_region      23362. .23462
/rpt_family="MIR"
repeat_region      complement(24063. .24094)
/rpt_family="GC-rich"

```

repeat\_region 24283..24582  
 /rpl\_family="Alusx"  
 repeat\_region 24583..24608  
 /rpl\_family="(GAAA)n"  
 repeat\_region 24851..24930  
 /rpl\_family="AluV/FLAM"  
 repeat\_region complement(24932..25220)  
 /rpl\_family="Alusx"  
 repeat\_region 25221..25284  
 /rpl\_family="AluV/FLAM"

## Alignment Scores:

Pred. No.: 3.29e-06 Length: 41006  
 Score: 90.00 Matches: 17  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-397-967-15 (1-17) x AC007201 (1-41006)

OY 1 AlalysleuProleuAspLyAspTyTyValValArgGluProGly 17

Db 11061 GCTAAGCTGCTGCCGCTTGACAAAGACTACTACTGCTGCCGACGACGAC 11011

## RESULT 12

MUSJAK3A 3528 bp mRNA linear ROD 24-APR-1996  
 LOCUS Mus musculus JAK3 gene, complete cds.  
 DEFINITION L40172.1 GI:1019909  
 ACCESSION JAK3 gene.  
 VERSION Mus musculus (strain BALB/c x 129 F2) (clone: clone-s12-18) neonate  
 SOURCE thymus CDNA to mRNA.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 3528)  
 Gurtlak,C.B. and Berg,L.J.  
 TITLE Murine JAK3 is preferentially expressed in hematopoietic tissues  
 and lymphocyte precursor cells  
 JOURNAL Blood 87 (8), 3151-3160 (1996)  
 MEDLINE 96184772  
 PUBMED 8605329

## FEATURES

source Location/Qualifiers  
 1..3528  
 /organism="Mus musculus"  
 /strain="BALB/c x 129 F2"  
 /db\_xref="taxon:10090"  
 /clone="clone-s12-18"  
 /cell\_type="T cell precursor"  
 /tissue\_type="thymus"  
 /dev\_stage="neonate"  
 1..3528  
 /gene="JAK3"  
 1..71  
 /gene="JAK3"  
 72..3374  
 /gene="JAK3"  
 /codon\_start=1  
 /protein\_id="AAC42085.1"  
 /db\_xref="GI:1019910"  
 /translation="MAPPEETPLIPORSCSLSSSEAGALHVLPPRGCPPORLIFS  
 EGDYLAEDLCVRKAKAGCILPVYHSLFALATEDPSCWPFPSHFIEDVDVLYRL  
 RYFPDWFGLTCHRGKRDLSALDLHLVLEFAOHRSDLVSGRLPVGSLMKRGG  
 EFLSLAVLDLQMAREQARQGLKLTYSYKALPLPSLRDVIOGQNFVRRTIRRYV  
 LALRRVACQADRALMAKYTLDLERLHPAATETETRGVPGQEPGGLRRAAGDGI  
 SWSSQDELFOFCDPEIVDVSIKQAPRGAGEHRLVTRMDGHILEAFPGIPE  
 PCGYILRSQDDYDFLIRACVTPPLGPDYKGLIQDPSGAFSLVGLSQPHRSIREL  
 LAACWMSGLVDAALNLVSCAPRKREKSNLIVRRCCTPAPAGCSPSCCALQOLS  
 FHTIPDLSLMEHNLGHGSPFKIFGRREVVYDGEFHDESEVLKLVMDSRHRMSFLL  
 EASLMSOVSYPHLVLLHGVCMAGDSIMVQEFVYLGALIDMYLKRKHLYSASMKLQVIT

3'UTR  
 polyA\_signal  
 BASE COUNT 648 a 1104 c 1036 g 740 t  
 ORIGIN  
 KOLAYALNLEDEKGLPHGNVSAKRVLLAREGDCGNPPRIKISDPEVSTYLSLEMLTD  
 RIPWVAPEIOAEOTLCLEADKMGKATTEGFEYSSGPAHITSLSPAKLKRYEPOGOL  
 PALKMTLGLITQCMAYDPGRPSFRAILDLNGLITSDYELLSDPGIPSPRDEL  
 CGCAQILYACODPAITFEERHLKITSILGNGNGSVELCYRDLGNTGPIVAVKOLDS  
 GPDQORDFORREIQILKALHSDFIYKRGVSGPGOSRLVMEYLPSCGLDFLORR  
 ARLHHDRLLEFAMQICKGMEYLGARCVHRLAARNITLVESEAHVKTADRCATLPL  
 GNDYVVERPGSPIFENYAPESLSDNIFSRQSDVMSRGVLYELFTYCDKSCSAER  
 LRMGPEREGPILCRLLBLAEGRLPRPPICPIEVQELMQLCWAPSDHDPARFTLS  
 FOLDALMRGPG"

## Alignment Scores:

Pred. No.: 5.63e-06 Length: 3528  
 Score: 83.00 Matches: 16  
 Percent Similarity: 94.12% Conservative: 0  
 Best Local Similarity: 94.12% Mismatches: 1  
 Query Match: 92.22% Indels: 0  
 DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x MUSJAK3A (1-3528)

OY 1 AlalysleuProleuAspLyAspTyTyValValArgGluProGly 17

Db 2970 GCTAAGCTGCTGCCCTTGGAAGACACTACTGCTGCCGACGCTGGC 3020

## RESULT 13

MUSPYKIN 3723 bp mRNA linear ROD 21-SEP-1994  
 LOCUS Mouse protein tyrosine kinase (JAK3) mRNA, complete cds.  
 DEFINITION L32955  
 ACCESSION L32955.1 GI:529238  
 VERSION Interleukin 2; Interleukin 4; Janus kinase; protein tyrosine  
 KEYWORDS kinase.  
 SOURCE Mus musculus (strain BALB/c, sub-species domesticus) CDNA to mRNA.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 3723)  
 Witthuhn,B.A., Silvennoinen,O., Miura,O., Lal,K.S., Cwik,C.,  
 Liu,E.T. and Ihle,J.N.  
 TITLE Involvement of the Jak-3 Janus kinase in signalling by Interleukins  
 2 and 4 in lymphoid and myeloid cells  
 JOURNAL Nature 370 (6485), 153-157 (1994)  
 MEDLINE 94294024  
 PUBMED 8022486  
 COMMENT This entry has been reported under the accession number L329555 in  
 the Nature article.  
 FEATURES  
 source Location/Qualifiers  
 1..3723  
 /organism="Mus musculus"  
 /strain="BALB/c"  
 /sub\_species="domesticus"  
 /db\_xref="taxon:10090"  
 /cell\_line="Abelson transformed pre-B-cell line 22D6"  
 1..3723  
 /cell\_type="lymphoid and myeloid"  
 /gene="jak3"  
 1..12  
 /gene="jak3"  
 13..3312  
 /gene="jak3"  
 /standard\_name="Janus kinase, mjak3"  
 /codon\_start=1  
 /evidence="experimental"  
 /product="protein tyrosine kinase"  
 /protein\_id="AA021565.1"  
 /db\_xref="GI:529239"  
 /translation="MAPPEETPLIPORSCSLSSSEAGALHVLPPRGCPPORLIFS

FGDYLAEDICVBARAKACALIPYHSHLAFALATEDSCWPPSHIFCIEDVDIOVLVYRL  
REFEFDGELTCHRFGLKRDLSALIDLVLEHLEFAOHRSDVSGRLPVGSMKEOG  
EFLSLAVLADLAQMARAOAPRPGELTKTVSKACLPSPEDVIOGONFVTRRRIRRTVY  
LALLPGCRLPGRPVLAAMARY11DLERLHPAATTEFRVCLPGAQEPGLLRACANGI  
PMSNNELFQTECFDEPEIIVDSINQAPRVGPGHEHRLVITVMDGHIIEAFPGIPEA  
LSVVALVDGFRILICDSRHCFKEVAPRPLEEADVCHGPTTLDFAIHKLAAASLP  
GVYILRRSDODYSFLLTACVOTPLGPDYKGLIRDDPSGAFSLVGLSQPHRSLELL  
AACNSGIRDAALYTSCCAPRKEKSNLIYVRGCDNAPAPGCSPECALTOLEF  
HTPTPSLHNEHLGHGSPFKIPRGSRRVGEETHSEVLLKMDSRHNCMESFLE  
AASLMGVSFPHILVYLHAGOMAGDSIMNOEPVYLAIDMYLRRGHVLSASKLOYTK  
OLAVYALNYLEDKLPHEVNSARKVILARREGDGNPPFIKISDPGSPVYLSLEMLDR  
IPWAPPECLDEAOTILEADKGFATWEPFORGAHITSLEPAKXIKFYEDCOOLP  
ALMWTETAGLITOCMAVDPGRPSFPAILLDLNGLITSDLELSDPTTPOIPSPRDLIC  
VAGQIYACODPAIFEEERHLKXISLKGKNGFSGELCRDPLDNTGPLVAVKOLDS  
VPPQDRDFOREIOILKALHSDIYKRYGVSGPQSLRLVMEYLPSGLDLLOLRH  
GLHTRDLRLPAMQICKMEYLGARCVGRDLARNLIVSEAHVKIADGGLALDPLG  
KDYVVERPGOSPFMWYAPRSLSDNIFRSQSDVMSFGVYVLELFTYCDSCSPSAEFL  
RMGGEREPPLCRLELLAEGRRLLPPTCPTEVOELMQLCWAPEPHDRPAPATLSP  
QLDPLMRGRPG"  
mat\_peptide  
/gene="jak3"  
/product="protein tyrosine kinase"  
/standard\_name="Janus Kinase, mujak3"  
/evidence=experimental  
3'UTR  
/gene="jak3"  
3723  
polyA\_site  
/gene="jak3"  
711 a 1162 c 1079 g 771 t  
BASE COUNT  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.96e-06 Length: 3723  
Score: 83.00 Matches: 16  
Percent Similarity: 94.12% Conservative: 0  
Best Local Similarity: 94.12% Mismatches: 1  
Query Match: 92.22% Indels: 0  
Gaps: 0  
DB: 10  
US-09-397-967-15 (1-17) x MUSPTYKIN (1-3723)  
OY 1 AlalysleuleuProleuasplysAsPTYrTyValaValArgGluProGly 17  
|||||  
Db. 2908 GCTAAGCTGCTGCCCTCGGAAGACTACTAGTGTCCGCGAGCCTGGC 2958  
RESULT 14  
MUSJAK3H 4016 bp mRNA linear ROD 13-SEP-1994  
LOCUS  
DEFINITION Mus musculus (clone 32D5-1) protein tyrosine kinase (JAK3) mRNA.  
ACCESSION L33768  
VERSION L33768.1 GI:508533  
KEYWORDS protein tyrosine kinase.  
SOURCE Mus musculus cDNA to mRNA.  
ORGANISM Mus musculus. Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS Rane,S.G. and Reddy,E.P.  
TITLE JAK3: a novel JAK kinase associated with terminal differentiation  
of hematopoietic cells  
JOURNAL Oncogene 9 (8), 2415-2423 (1994)  
MEDLINE 94309920  
PUBMED 7518579  
FEATURES  
source  
1..4016  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="32D5-1"  
/cell\_line="32Dc13"  
/tissue\_lib="lambda gtl11"  
7..3906  
/note="protein tyrosine kinase"  
/codon\_start=1

/product="JAK3"  
/protein\_id="AAA21415.1"  
/db\_xref="GI:37607"  
/translation="MAPPSEPTPLIPORSCSLSSSEAGALHVLPPRGQPPORLSFS  
FGDYLAEDICVBARAKACALIPYHSHLAFALATEDSCWPPSHIFCIEDVDIOVLVYRL  
EFLSLAVLADLAQMARAOAPRPGELTKTVSKACLPSPEDVIOGONFVTRRRIRRTVY  
LALLPGCRLPGRPVLAAMARY11DLERLHPAATTEFRVCLPGAQEPGLLRACANGI  
PMSNNELFQTECFDEPEIIVDSINQAPRVGPGHEHRLVITVMDGHIIEAFPGIPEA  
LSVVALVDGFRILICDSRHCFKEVAPRPLEEADVCHGPTTLDFAIHKLAAASLP  
GVYILRRSDODYSFLLTACVOTPLGPDYKGLIRDDPSGAFSLVGLSQPHRSLELL  
AACNSGIRDAALYTSCCAPRKEKSNLIYVRGCDNAPAPGCSPECALTOLEF  
HTPTPSLHNEHLGHGSPFKIPRGSRRVGEETHSEVLLKMDSRHNCMESFLE  
AASLMGVSFPHILVYLHAGOMAGDSIMNOEPVYLAIDMYLRRGHVLSASKLOYTK  
OLAVYALNYLEDKLPHEVNSARKVILARREGDGNPPFIKISDPGSPVYLSLEMLDR  
IPWAPPECLDEAOTILEADKGFATWEPFORGAHITSLEPAKXIKFYEDCOOLP  
ALMWTETAGLITOCMAVDPGRPSFPAILLDLNGLITSDLELSDPTTPOIPSPRDLIC  
VAGQIYACODPAIFEEERHLKXISLKGKNGFSGELCRDPLDNTGPLVAVKOLDS  
VPPQDRDFOREIOILKALHSDIYKRYGVSGPQSLRLVMEYLPSGLDLLOLRH  
GLHTRDLRLPAMQICKMEYLGARCVGRDLARNLIVSEAHVKIADGGLALDPLG  
KDYVVERPGOSPFMWYAPRSLSDNIFRSQSDVMSFGVYVLELFTYCDSCSPSAEFL  
RMGGEREPPLCRLELLAEGRRLLPPTCPTEVOELMQLCWAPEPHDRPAPATLSP  
QLDPLMRGRPG"  
misc\_feature  
/note="JAK homology domain II (JH2)"  
2199..3024  
misc\_feature  
/note="JAK homology domain I (JH1)"  
3076..3876  
BASE COUNT 742 a 1203 c 1235 g 836 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.47e-06 Length: 4016  
Score: 83.00 Matches: 16  
Percent Similarity: 94.12% Conservative: 0  
Best Local Similarity: 94.12% Mismatches: 1  
Query Match: 92.22% Indels: 0  
Gaps: 0  
DB: 10  
US-09-397-967-15 (1-17) x MUSJAK3H (1-4016)  
OY 1 AlalysleuleuProleuasplysAsPTYrTyValaValArgGluProGly 17  
|||||  
Db. 3544 GCTAAGCTGCTGCCCTCGGAAGACTACTAGTGTCCGCGAGCCTGGC 3594  
RESULT 15  
MUSJAK3H 10326 bp DNA linear ROD 15-JAN-1997  
LOCUS  
DEFINITION Mus musculus protein tyrosine kinase JAK3 gene, complete cds.  
ACCESSION U71201  
VERSION U71201.1 GI:1178748  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus. Chordata; Craniata; Vertebrata; Euteleostomi;  
Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS Kumar,A., Toscani,A., Rane,S. and Reddy,E.P.  
TITLE Structural organization and chromosomal mapping of JAK3 locus  
JOURNAL Oncogene 13 (9), 2009-2014 (1996)  
MEDLINE 8934548  
PUBMED 8934548  
REFERENCE  
AUTHORS Kumar,A. and Reddy,E.P.  
TITLE Direct Submission  
JOURNAL Submitted (18-SEP-1996) Fels Institute, Temple University, 3307 N.  
Broad St, Rm 339 AHB, Philadelphia, PA 19140, USA  
FEATURES  
source  
1..10326  
/organism="Mus musculus"  
/strain="129/J"

/db\_xref="taxon:10090"  
 /chromosome="19"  
 /map="19p13.1-p13.2"

MRNA  
 join(92, 287, 441, .564, 740, .851, 963, .1108, 1317, .1599,  
 2170, .2289, 2372, .2529, 2617, .2728, 3230, .3413, 3781, .3902,  
 3998, .4129, 4596, .4680, 5022, .5149, 5642, .5774, 5905, .6056,  
 6177, .6327, 6856, .6995, 7074, .7260, 7427, .7551, 7629, .7801,  
 8524, .8641, 8907, .9017, 9707, .10212)  
 /product="JAK3"

CDS  
 join(104, .287, 441, .564, 740, .851, 963, .1108, 1317, .1599,  
 2170, .2289, 2372, .2529, 2617, .2728, 3230, .3413, 3781, .3902,  
 3998, .4129, 4596, .4680, 5022, .5149, 5642, .5774, 5905, .6056,  
 6177, .6327, 6856, .6995, 7074, .7260, 7427, .7551, 7629, .7801,  
 8524, .8641, 8907, .9017, 9707, .9772)  
 /function="protein tyrosine kinase"  
 /codon\_start=1  
 /product="JAK3"  
 /protein\_id="AAB40917.1"  
 /db\_xref="GI:1778749"

/translation="MAPSEETPLIPQSCSLSSSEAGALHVLPPRGPGPQRLSFS  
 FGDIADLCVRAKACGILPVYHSLFALATEDSCWEPSSHIFCIEDVDTQVLYRL  
 RFYEPDWFGLETCHRFGLRKDLTSAILDLHLFLFAOHRSDLVSGRLPYGLSMKEOG  
 EFLSLAVLDLAOMAREQAOORGEILKTVSKACLPISBDVIOGNFVTRRIKRTIV  
 LALRWSPARPTAHGOVYLDLELHPATTTTFRVGLPQAEPEGLLAVAGDNGIS  
 WSSGDOELFQTFQDFPRIVVSTIKOPTCGSGREHRLVTVTRMDGHILEAFPGLEAL  
 SPVALNDGVRFLICDSRHRYCKEVAAPRLLEEAEALHGPIITLDFAIHRKLAASLPG  
 TYILRNSPDYDSFLTACVOTPLGPYKGLIROPDSGAFSLVGSPTACGTCLO  
 CMNSGLVDDAALNLTSCCAPRPEKSNLIVRRCTPAPAPCSPSCCALIQLSFTH  
 IPIDSLGHEMLHGGSFTKIFGRRRREVDETHDESEVLKVMDSRRHNCMESLEAS  
 LMSQVSPHLVLHGVCMAGDSIMVOEFYIGAIDMYLRKGHVSAWKLOYTKOLA  
 YALNYLEDKGLPHGNVARKVILLAREGGDGNPFITLSDPGVSPYLSLEMLIDRIPLW  
 VAPECLQEAQTLCTLEADKNGFGATTWVEVSGCPAHITSLLEPAKKLFYEDQQLPALK  
 WTELAGLITQCMAYDPGRNRSFRATLRDLNGLITSDYELLSDPTGIPSPDELCDGA  
 QIYACODPAIFEEHRLKYISLGKNGFVSELCRYDPLDNTGPLVAVKOLQHSQPDQ  
 RDRREIQIILKALHSDFIKRVGVSYPGROSILVMEYLPSCGLRDFLQHRALHT  
 DRLLFAWQICKMEYLGARCVHRLAARNTLVESEAHVKIADFGIAKLILPIGKDY  
 VVERPGOSPFWAPESLSDNIFSRQSDVSGVYVLELFTYCDKSCSPSAEFLMMG  
 PEREGPPLCLLELLAEGRRLPPPTCPTEVOELMOCVAPSRITGQPSAP"

BASE COUNT 2088 a 3007 c 2880 g 2300 t 51 others

## Alignment Scores:

Prod. No.:	1.78e-05	Length:	10326
Score:	83.00	Matches:	16
Percent Similarity:	94.12%	Conservative:	0
Best Local Similarity:	94.12%	Mismatches:	1
Query Match:	92.22%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-15 (1-17) x MMU71201 (1-10326)

QY 1 AlalysleuDeupProleuAspTYTYVaiValArgLupProGly 17  
 DB 7734 GCTAAGCTGCTGCCCTGGGAAAGACTACTAGCTGCTCCGAGAGCTGGC 7784

Search completed: April 28, 2003, 15:56:39  
 Job time : 114.732 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: April 28, 2003, 13:51:29 ; Search time 2.33065 Seconds  
(without alignments)  
2236.936 Million cell updates/sec

Title: US-09-397-967-15  
Perfect score: 90  
Sequence: 1 AKLLPLDKDYVREPG 17

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=xlp  
-O=/cgn2\_1/USPRO.spool/US09397967/cunat\_26042003.182313.10750/app.query.fasta\_1.1486  
-DB=Issued\_Patents.NA -GPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967 -CGN\_1\_1\_59.etrnat.26042003.182313.10750 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=130 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued\_Patents.NA:\*

1: /cgn2\_6/ptodata/2/1na/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/1na/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/1na/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/1na/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/1na/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/1na/backfilest.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	3807	1	US-08-357-598-1
2	90	100.0	3807	2	US-09-003-289-1
3	90	100.0	3807	5	PCT-US95-16435-1
4	83	92.2	4016	5	PCT-US95-08354A-1
5	64	71.1	3435	4	US-09-046-158A-21
6	64	71.1	3495	1	US-08-446-038B-2
7	64	71.1	3495	1	US-08-446-010B-2
8	64	71.1	3495	1	US-08-805-445-2
9	64	71.1	3495	2	US-08-064-067D-2
10	64	71.1	3495	2	US-09-066-208-2
11	64	71.1	3495	4	US-08-980-080-3
12	64	71.1	3629	1	US-08-097-997A-8

13	64	71.1	3629	3	US-08-665-574C-8	Sequence 8, Appl1
14	64	71.1	3629	4	US-08-946-994-8	Sequence 8, Appl1
15	64	71.1	4482	2	US-08-567-508C-1	Sequence 1, Appl1
16	64	71.1	4482	3	US-09-146-480-1	Sequence 1, Appl1
17	46	51.1	1863	2	US-08-455-073A-3	Sequence 3, Appl1
18	44	48.9	3147	4	US-09-101-886B-1	Sequence 3, Appl1
19	44	48.9	11703	4	US-09-101-886B-3	Sequence 3, Appl1
20	43	47.8	978	4	US-09-134-001C-2696	Sequence 12, Appl
21	43	47.8	3561	3	US-08-097-997A-12	Sequence 12, Appl
22	43	47.8	3561	3	US-08-665-574C-12	Sequence 12, Appl
23	43	47.8	3561	4	US-08-946-994-12	Sequence 12, Appl
24	43	47.8	4052	2	US-08-833-228-1	Sequence 1, Appl1
25	43	47.8	5024	1	US-08-920-812-7	Sequence 7, Appl1
26	43	47.8	5024	1	US-08-920-827-7	Sequence 7, Appl1
27	43	47.8	5024	1	US-08-921-177-7	Sequence 7, Appl1
28	43	47.8	5024	1	US-08-362-577C-7	Sequence 7, Appl1
29	43	47.8	5024	2	US-08-920-828-7	Sequence 7, Appl1
30	42	46.7	2734	3	US-09-135-021-79	Sequence 7, Appl1
31	42	46.7	2821	4	US-09-135-010A-115	Sequence 7, Appl
32	42	46.7	2821	4	US-09-597-735-115	Sequence 115, App
33	42	46.7	2821	4	US-09-597-735-115	Sequence 115, App
34	42	46.7	3181	3	US-09-135-021-1	Sequence 1, Appl1
35	42	46.7	3181	4	US-09-135-020-1	Sequence 1, Appl1
36	42	46.7	3181	4	US-09-135-010A-1	Sequence 1, Appl1
37	42	46.7	3181	4	US-09-444-871-1	Sequence 1, Appl1
38	42	46.7	3181	4	US-09-597-735-1	Sequence 1, Appl1
39	42	46.7	3181	4	US-09-444-295-1	Sequence 1, Appl1
40	42	46.7	3181	4	US-09-597-732-1	Sequence 1, Appl1
41	42	46.7	3181	3	US-09-135-021-5	Sequence 5, Appl1
42	41	45.6	998	4	US-09-540-014-26	Sequence 26, Appl
43	41	45.6	3355	2	US-08-933-821-3	Sequence 3, Appl1
44	41	45.6	3355	3	US-08-960-507-3	Sequence 3, Appl1
45	41	45.6	3355	4	US-09-136-828-3	Sequence 3, Appl1

## ALIGNMENTS

RESULT 1  
US-08-357-598-1  
Sequence 1, Application US/08357598  
Patent No. 5705625  
GENERAL INFORMATION:  
APPLICANT: Civin, Curt I.  
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/357,598  
FILING DATE: 15-DEC-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/033001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3807 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-357-598-1

Alignment Scores:  
Pred. No.: 2.23e-07 Length: 3807  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-15 (1-17) x US-08-357-598-1 (1-3807)

OY 1 AlalysleuLeuProLeuAspLysAspTyrTyrrValValaArgGluProGly 17  
Db 3069 GCTAAGCTGCTGCCGCTTGACAAAGACTACTACGTGGTCCGCGACCGCAGGC 3119

RESULT 2  
US-09-003-289-1  
Sequence 1, Application US/09003289  
Patent No. 5916792  
GENERAL INFORMATION:  
APPLICANT: Civin, Curt I.  
APPLICANT: Small, Donald  
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/003,289  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/357,598  
FILING DATE: 15-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/033001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3807 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-09-003-289-1

Alignment Scores:  
Pred. No.: 2.23e-07 Length: 3807  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-09-397-967-15 (1-17) x US-09-003-289-1 (1-3807)

OY 1 AlalysleuLeuProLeuAspLysAspTyrTyrrValValaArgGluProGly 17  
Db 3069 GCTAAGCTGCTGCCGCTTGACAAAGACTACTACGTGGTCCGCGACCGCAGGC 3119

RESULT 3  
PCT-US95-16435-1  
Sequence 1, Application PC/TUS9516435  
GENERAL INFORMATION:  
APPLICANT: The Johns Hopkins University School of Medicine  
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/16435  
FILING DATE: 15-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Halle, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/033W01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5099  
TELEFAX: 619/678-5070  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3807 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
PCT-US95-16435-1

Alignment Scores:  
Pred. No.: 2.23e-07 Length: 3807  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-09-397-967-15 (1-17) x PCT-US95-16435-1 (1-3807)

OY 1 AlalysleuLeuProLeuAspLysAspTyrTyrrValValaArgGluProGly 17  
Db 3069 GCTAAGCTGCTGCCGCTTGACAAAGACTACTACGTGGTCCGCGACCGCAGGC 3119

RESULT 4  
PCT-US95-08354A-1  
Sequence 1, Application PC/TUS9508354A  
GENERAL INFORMATION:  
APPLICANT: Temple University - Of The  
APPLICANT: Commonwealth System of Higher Education  
TITLE OF INVENTION: JAK3 PROTEIN TYROSINE  
TITLE OF INVENTION: KINASE AND DNA ENCODING THE SAME  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seidel, Gonda, Lavorina  
ADDRESS: 6 Monaco, P.C.  
STREET: Suite 1800, Two Penn Center  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.



ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 Inch, 720 Kb  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT-US95/08354A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/272,368  
FILING DATE: 8 July 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 6056-203 PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4016 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: single stranded  
TOPOLOGY: linear  
PCT-US95-08354A-1

Alignment Scores:  
Pred. No.: 4.97e-06 Length: 4016  
Score: 83.00 Matches: 16  
Percent Similarity: 94.12% Conservative: 0  
Best Local Similarity: 94.12% Mismatches: 1  
Query Match: 92.22% Indels: 0  
DB: 5 Gaps: 0

US-09-397-967-15 (1-17) x PCT-US95-08354A-1 (1-4016)

QY 1 A1AlaLysLeuLeuProLeuAspTyrTyrValValArgGluProGly 17  
DB 3544 GCTAAGCTGCTGCCCTGGCAAGGACTACTGCTGCCGCGACCTGTC 3594

RESULT 5  
US-09-046-158A-21  
Sequence 21, Application US/09046158A  
Patent No. 6187552  
GENERAL INFORMATION:  
APPLICANT: Roberts, Steven L.  
TITLE OF INVENTION: METHOD FOR IDENTIFYING INHIBITORS OF  
JAK2/CYTOKINE RECEPTOR BINDING  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pharmacia & Upjohn Co., Intellectual Property  
ADDRESS: Legal Services  
STREET: 301 Henrietta Street  
CITY: Kalamazoo  
STATE: MI  
COUNTRY: USA  
ZIP: 49001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/046,158A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Darnley Jr., James D.  
REGISTRATION NUMBER: 33,673  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 616/833-2210  
TELEFAX: 616/833-8897  
TELEX: 224401  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3435 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-046-158A-21

Alignment Scores:  
Pred. No.: 0.0159 Length: 3435  
Score: 64.00 Matches: 11  
Percent Similarity: 87.50% Conservative: 3  
Best Local Similarity: 68.75% Mismatches: 2  
Query Match: 71.11% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967-15 (1-17) x US-09-046-158A-21 (1-3435)

QY 2 LysLeuLeuProLeuAspTyrTyrValValArgGluProGly 17  
DB 2998 AAAGTCTTCCACACAGCAAGATCTATTAAGTAAAGAACTGCT 3045

RESULT 6  
US-08-446-038B-2  
Sequence 2, Application US/08446038B  
Patent No. 5658791  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;  
APPLICANT: Harpur, Alisa  
TITLE OF INVENTION: No. 5658791el Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felle & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,038B  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/064,067  
FILING DATE: 30-Jun-1993  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian 88229/91  
FILING DATE: 27-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5658791man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5244  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-838-3884  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3495 base pairs  
TYPE: nucleic acid

```
STRANDEDNESS: single
;
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acid
US-08-446-038B-2

Alignment Scores:
Pred. No.: 0.0162 Length: 3495
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-15 (1-17) x US-08-446-038B-2 (1-3495)
Qy 2 LysleuProLeuAspLysAspTyrTyrValValArgGluProGly 17
Db 2587 AAAGCTTCCCGCAGACAAAGATCTACAAAGTAAAGACCGAGG 2634

RESULT 7
US-08-446-010B-2
; Sequence 2, Application US/08446010B
; Patent No. 5716818
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiacki, Andrew;
; APPLICANT: Harpur, Alisa
; TITLE OF INVENTION: NO. 5716818el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,010B
; FILING DATE: 19-May-1995
; CLASSIFICATION: 433
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/446,038
; FILING DATE: 19-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/064,067
; FILING DATE: 30-Jun-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5716818-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5716818-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian 88229/91
; FILING DATE: 27-No. 5716818-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Baer, Madeline F.
; REGISTRATION NUMBER: 36,437
; REFERENCE/DOCKET NUMBER: LUD 5244.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acid
```

```
US-08-446-010B-2

Alignment Scores:
Pred. No.: 0.0162 Length: 3495
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-15 (1-17) x US-08-446-010B-2 (1-3495)
Qy 2 LysleuProLeuAspLysAspTyrTyrValValArgGluProGly 17
Db 2587 AAAGCTTCCCGCAGACAAAGATCTACAAAGTAAAGACCGAGG 2634

RESULT 8
US-08-805-445-2
; Sequence 2, Application US/08805445
; Patent No. 5821069
; GENERAL INFORMATION:
; APPLICANT: Wilks, Andrew F.; Ziemiacki, Andrew;
; APPLICANT: Harpur, Alisa
; TITLE OF INVENTION: NO. 5821069el Protein Tyrosine Kinase
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/805,445
; FILING DATE: 25-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,038
; FILING DATE: 19-MAY-1995
; APPLICATION NUMBER: 08/064,067
; FILING DATE: 30-Jun-1993
; APPLICATION NUMBER: PCT/US91/08889
; FILING DATE: 26-No. 5821069-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian PK3594/90
; FILING DATE: 28-No. 5821069-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Australian 88229/91
; FILING DATE: 27-No. 5821069-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5821069man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5244
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-688-9200
; TELEFAX: 212-838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acid
US-08-805-445-2

Alignment Scores:
Pred. No.: 0.0162 Length: 3495
Score: 64.00 Matches: 11
```

Percent Similarity: 87.504  
Best Local Similarity: 68.758  
Query Match: 71.11%  
DB: 1  
Indels: 0  
Gaps: 0

US-09-397-967-15 (1-17) x US-08-805-445-2 (1-3495)

OY 2 LysleuleuProleuAsplysAspyrTyTyValValArgGluProgly 17  
DB 2587 AAGGCTTGCCGACGACAAAGAAATACTACAAAGTAAGAGACCGACGG 2634

## RESULT 9

US-08-064-067D-2  
Sequence 2, Application US/08064067D  
Patent No. 5852184  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;  
APPLICANT: Harpur, Alisa  
TITLE OF INVENTION: No. 5852184el Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felte & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/064, 067D  
FILING DATE: 30-Jun-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-No. 5852184-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-No. 5852184-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian 88229/91  
FILING DATE: 27-No. 5852184-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5852184man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5244  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-838-3884  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
US-08-064-067D-2

Alignment Scores:  
Pred. No.: 0.0162 Length: 3495  
Score: 64.00 Matches: 11  
Percent Similarity: 87.508  
Best Local Similarity: 68.758  
Query Match: 71.11%  
DB: 2  
Indels: 0  
Gaps: 0

US-09-397-967-15 (1-17) x US-08-064-067D-2 (1-3495)

OY 2 LysleuleuProleuAsplysAspyrTyTyValValArgGluProgly 17  
DB 2587 AAGGCTTGCCGACGACAAAGAAATACTACAAAGTAAGAGACCGACGG 2634

## RESULT 10

US-09-066-208-2  
Sequence 2, Application US/09066208  
Patent No. 5910426  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;  
APPLICANT: Harpur, Alisa  
TITLE OF INVENTION: No. 5910426el Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felte & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/066, 208  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/805, 445  
FILING DATE: 25-FEB-1997  
APPLICATION NUMBER: US 08/446, 038  
FILING DATE: 19-MAY-1995  
APPLICATION NUMBER: 08/064, 067  
FILING DATE: 30-Jun-1993  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-No. 5910426-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-No. 5910426-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian 88229/91  
FILING DATE: 27-No. 5910426-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5910426man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5244  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-838-3884  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
US-09-066-208-2

Alignment Scores:  
Pred. No.: 0.0162 Length: 3495  
Score: 64.00 Matches: 11  
Percent Similarity: 87.508  
Best Local Similarity: 68.758  
Query Match: 71.11%  
DB: 2  
Indels: 0  
Gaps: 0

US-09-397-967-15 (1-17) x US-09-066-208-2 (1-3495)

OY 2 LysleuleuProleuAsplysAspyrTyTyValValArgGluProgly 17  
DB 2587 AAGGCTTGCCGACGACAAAGAAATACTACAAAGTAAGAGACCGACGG 2634  
RESULT 11  
US-08-980-080-3

Sequence 3, Application US/08980080  
Patent No. 6312941  
GENERAL INFORMATION:  
APPLICANT: CARTER-SU, CHRISTIN  
APPLICANT: RUI, LIANG-YOU  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING  
TITLE OF INVENTION: SIGNALING PATHWAY AGONISTS AND ANTAGONISTS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MEDLEN & CARROLL, LLP  
STREET: 220 MONTGOMERY STREET, SUITE 2200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/980,080  
FILING DATE: 26-NOV-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CARROLL, PETER G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: UM-03036  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2979  
US-08-980-080-3

Alignment Scores:  
Pred. No.: 0.0162 Length: 3495  
Score: 64.00 Matches: 11  
Percent Similarity: 87.50% Conservative: 3  
Best Local Similarity: 68.75% Mismatches: 2  
Query Match: 71.11% Indels: 0  
Gaps: 0  
DB: 4

US-09-397-967-15 (1-17) x US-08-980-080-3 (1-3495)

QY 2 LysleuleuproleuasplysAsPTyTyValValArgGUpProgly 17  
DB 2587 AAAGTCTTCCGCGACGACAAAGTAATCTACAAAGTAAGGCGCCAGGG 2634  
||||:||||| |||||:||||| |||||:||||| |||||:||||| |||||:|||||

RESULT 12  
US-08-097-997A-8  
Sequence 8, Application US/08097997A  
Patent No. 5728536  
GENERAL INFORMATION:  
APPLICANT: Ihle, James N.  
APPLICANT: Silvenoinen, Olli  
APPLICANT: Wiltuhn, Bruce A.  
APPLICANT: Quelle, Frederick W.  
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal  
TITLE OF INVENTION: Transduction  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
City: Washington

STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,997A  
FILING DATE: 29-JULY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0370000/SLE/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3629 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 94..3480  
US-08-097-997A-8

Alignment Scores:  
Pred. No.: 0.0169 Length: 3629  
Score: 64.00 Matches: 11  
Percent Similarity: 87.50% Conservative: 3  
Best Local Similarity: 68.75% Mismatches: 2  
Query Match: 71.11% Indels: 0  
Gaps: 0  
DB: 1

US-09-397-967-15 (1-17) x US-08-097-997A-8 (1-3629)

QY 2 LysleuleuproleuasplysAsPTyTyValValArgGUpProgly 17  
DB 3088 AAAGTCTTCCGCGACGACAAAGTAATCTACAAAGTAAGGCGCCAGGG 3135  
||||:||||| |||||:||||| |||||:||||| |||||:||||| |||||:|||||

RESULT 13  
US-08-665-574C-8  
Sequence 8, Application US/08665574C  
Patent No. 6136595  
GENERAL INFORMATION:  
APPLICANT: Ihle, James N.  
APPLICANT: Silvenoinen, Olli  
APPLICANT: Wiltuhn, Bruce A.  
TITLE OF INVENTION: Signal Transduction  
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,574C  
FILING DATE: 18-JUN-1996  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/282,012

```

1  APPLICATION NUMBER: 08/097,997
2  FILING DATE: 29-JUL-1993
3  PRIOR APPLICATION DATA:
4  APPLICATION NUMBER: 08/118,968
5  FILING DATE: 09-SEP-1993
6  ATTORNEY/AGENT INFORMATION:
7  NAME: Fox, Samuel L.
8  REGISTRATION NUMBER: 30,353
9  REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/GKT
10 TELECOMMUNICATION INFORMATION:
11 TELEPHONE: (202) 371-2600
12 TELEFAX: (202) 371-2540
13 INFORMATION FOR SEQ ID NO: 8:
14 SEQUENCE CHARACTERISTICS:
15 LENGTH: 3629 base pairs
16 TYPE: nucleic acid
17 STRANDEDNESS: single
18 TOPOLOGY: linear
19 FEATURE:
20 NAME/KEY: CDS
21 LOCATION: 94..3480
22 US-08-946-994-8
23
24 Alignment Scores:
25 Pred. No.: 0.0169 Length: 3629
26 Score: 64.00 Matches: 11
27 Percent Similarity: 87.50% Conservative: 3
28 Best Local Similarity: 68.75% Mismatches: 2
29 Query Match: 71.1% Indels: 0
30 DB: 4 Gaps: 0
31
32 US-09-397-967-15 (1-17) x US-08-946-994-8 (1-3629)
33
34 QY 2 LysLeuLeuProLeuAspLysAspTyrTyrValValArgLysProCly 17
35 Db 3088 AAGGCTTGCCGACGACAAAGATACTACAAAGTAAAGAGCAGGCG 3135S
36
37 RESULT 15
38 US-08-567-508C-1
39 Sequence 1, Application US/08567508C
40 Patent No. 5914393
41 GENERAL INFORMATION:
42 APPLICANT: Coleman, Roger
43 APPLICANT: Stuart, Susan G.
44 TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE
45 NUMBER OF SEQUENCES: 5
46 CORRESPONDENCE ADDRESS:
47 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
48 STREET: 3174 Porter Drive
49 CITY: Palo Alto
50 STATE: CA
51 COUNTRY: US
52 ZIP: 94304
53 COMPUTER READABLE FORM:
54 MEDIUM TYPE: Diskette
55 COMPUTER: IBM Compatible
56 OPERATING SYSTEM: DOS
57 SOFTWARE: FASTSEQ Version 1.5
58 CURRENT APPLICATION DATA:
59 APPLICATION NUMBER: US/08/567,508C
60 FILING DATE: 05-DEC-1995
61 CLASSIFICATION: 435
62 PRIOR APPLICATION DATA:
63 APPLICATION NUMBER:
64 FILING DATE:
65 ATTORNEY/AGENT INFORMATION:
66 NAME: Billings, Lucy J.
67 REGISTRATION NUMBER: 36,749
68 REFERENCE/DOCKET NUMBER: PF-004905
69 TELECOMMUNICATION INFORMATION:
70 TELEPHONE: 650-845-0555
71 TELEFAX: 650-845-4166
72 INFORMATION FOR SEQ ID NO: 1:

```

SEQUENCE CHARACTERISTICS:  
LENGTH: 4482 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
LIBRARY: Placenta  
CLONE: 179527  
US-08-567-508C-1

Alignment Scores:  
Pred. No.: 0.0217 Length: 4482  
Score: 64.00 Matches: 11  
Percent Similarity: 87.50% Conservative: 3  
Best Local Similarity: 68.75% Mismatches: 2  
Query Match: 71.11% Indels: 0  
DB: 2 Gaps: 0

US-09-397-967-15 (1-17) x US-08-567-508C-1 (1-4482)

QY 2 LysLeuLeuProLeuAspLysAspTyrTyrValValArgGluProGly 17  
||||:||||| |||||:||||| |||||:|||||  
DB 3334 AAGCTCTGCCACACAGACAAAGATACCTATAAGTAAAGAAGAACTGCT 3381

Search completed: April 28, 2003, 16:57:51  
Job time : 6.33065 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 28, 2003, 16:07:34 ; Search time 4.67652 Seconds  
(without alignments)  
3955.543 Million cell updates/sec

Title: US-09-397-967-15  
Perfect score: 90  
Sequence: 1 AKLPLDKDYVVRPG 17

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 709820 seqs, 544064369 residues

Total number of hits satisfying chosen parameters: 1419640

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ .p2n .model -DEV=xlp  
-O=/cg2\_1/USPRO.spool/US09397967/rnat\_26042003\_182314\_10830/app.query.fasta\_1.1486  
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=plowsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09397967-@CGN\_1\_1\_202\_@rnat\_26042003\_182314\_10830  
-NCPU=6 -ICPU=3 -NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT  
-DISPATCH=100 -LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10  
-XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA.\*  
1: /cg2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cg2\_6/ptodata/1/pubpna/PC1\_NEW\_PUB.seq.\*  
3: /cg2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cg2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cg2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cg2\_6/ptodata/1/pubpna/PC1US\_PUBCOMB.seq.\*  
7: /cg2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cg2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cg2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cg2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cg2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cg2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cg2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cg2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	48	53.3	32190	9	US-10-091-504-2209
2	48	53.3	32190	10	US-09-764-869-2209
3	45	50.0	212	10	US-09-983-965-1232
4	45	50.0	481	9	US-10-092-154-1382

C	5	45	50.0	481	10	US-09-764-847-1382	Sequence 1382, Ap
C	6	45	50.0	861	10	US-09-770-445-589	Sequence 589, Ap
C	7	45	50.0	1152	9	US-09-938-842A-2486	Sequence 2486, Ap
C	8	45	50.0	22786	10	US-09-764-877-3727	Sequence 3727, Ap
C	9	44	48.9	405	10	US-09-960-352-1663	Sequence 1663, Ap
C	10	44	48.9	411	10	US-09-960-352-11844	Sequence 11844, A
C	11	44	48.9	430	9	US-09-736-457-48	Sequence 48, Ap
C	12	44	48.9	430	9	US-09-902-941-48	Sequence 48, Ap
C	13	44	48.9	430	9	US-09-849-626-48	Sequence 48, Ap
C	14	44	48.9	430	9	US-10-017-754-48	Sequence 48, Ap
C	15	44	48.9	430	10	US-09-998-598-1085	Sequence 1085, Ap
C	16	44	48.9	579	9	US-10-060-036-2884	Sequence 2884, Ap
C	17	44	48.9	1387	9	US-09-925-302-359	Sequence 359, Ap
C	18	44	48.9	1638	7	US-08-976-063C-9	Sequence 38, Ap
C	19	44	48.9	2013	9	US-09-938-842A-1519	Sequence 1519, Ap
C	20	44	48.9	3931	10	US-09-880-107-3365	Sequence 3365, Ap
C	21	44	48.9	32679	7	US-08-976-063C-1	Sequence 1, Ap
C	22	43	47.8	259	9	US-09-925-922-3	Sequence 3, Ap
C	23	43	47.8	365	10	US-09-878-574-2001	Sequence 2001, Ap
C	24	43	47.8	372	10	US-09-864-761-3611	Sequence 3611, Ap
C	25	43	47.8	410	10	US-09-864-761-32588	Sequence 32588, A
C	26	43	47.8	481	10	US-09-864-761-20379	Sequence 16094, A
C	27	43	47.8	654	10	US-09-864-761-16094	Sequence 6, Ap
C	28	43	47.8	1149	9	US-10-132-812-5	Sequence 5, Ap
C	29	43	47.8	1155	9	US-10-132-812-3	Sequence 3, Ap
C	30	43	47.8	1155	10	US-09-764-556-1	Sequence 1, Ap
C	31	43	47.8	1158	9	US-10-132-812-1	Sequence 1, Ap
C	32	43	47.8	1170	9	US-09-925-922-1	Sequence 1, Ap
C	33	43	47.8	1287	9	US-09-510-333-102	Sequence 102, Ap
C	34	43	47.8	1387	10	US-09-750-373-13	Sequence 13, Ap
C	35	43	47.8	2811	9	US-09-938-842A-2538	Sequence 2538, Ap
C	36	43	47.8	4846	7	US-08-781-986A-276	Sequence 276, Ap
C	37	43	47.8	42999	10	US-09-740-029-3	Sequence 3, Ap
C	38	43	47.8	44848	9	US-09-988-113-42	Sequence 42, Ap
C	39	43	47.8	44848	10	US-09-776-874A-42	Sequence 4, Ap
C	40	42.5	47.2	49744	10	US-09-927-091-4	Sequence 1, Ap
C	41	42.5	47.2	414	9	US-09-918-995-17687	Sequence 17687, A
C	42	42	46.7	470	9	US-09-918-995-22620	Sequence 22620, A
C	43	42	46.7	483	9	US-09-918-995-13854	Sequence 13854, A
C	44	42	46.7	2821	10	US-09-880-107-3358	Sequence 3358, Ap
C	45	42	46.7	2821	10	US-09-880-107-3358	Sequence 3358, Ap

## ALIGNMENTS

RESULT 1  
US-10-091-504-2209  
: Sequence 2209, Application US/10091504  
: Publication No. US20030059908A1  
: GENERAL INFORMATION:  
: APPLICANT: Rosen et al.  
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
: FILE REFERENCE: PC007C1  
: CURRENT APPLICATION NUMBER: US/10/091,504  
: CURRENT FILING DATE: 2002-03-07  
: NUMBER OF SEQ ID NOS: 2442  
: PRIOR Application removed - See File Wrapper or Palm  
: SOFTWARE: Patencin Ver. 2.0  
: SEQ ID NO 2209  
: LENGTH: 32190  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
US-10-091-504-2209

Alignment Scores:  
Pred. No.: 268  
Score: 48.00  
Percent Similarity: 75.00%  
Best Local Similarity: 50.00%  
Query Match: 53.33%  
DB: 9  
Length: 32190  
Matches: 8  
Conservative: 4  
Mismatch: 4  
Indels: 0  
Gaps: 0  
US-09-397-967-15 (1-17) x US-10-091-504-2209 (1-32190)





```
US-09-770-445-589/C
; Sequence 589, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Mathew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, MaJa
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2023US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445.
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 589
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-445-589

Alignment Scores:
Pred. No.: 12.6 Length: 861
Score: 45.00 Matches: 9
Percent Similarity: 78.57% Conservative: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x US-09-770-445-589 (1-861)
QY 4 LeuProLeuAspLysAspTyrTyrValValArgGluProGly 17
DB 524 CTTGACCTTGATGAGATGTTATGTTGACCAAGCCAGGT 483

RESULT 7
US-09-938-842A-2486
; Sequence 2486, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepes, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: S100-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
```

```
; SEQ ID NO 2486
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2486

Alignment Scores:
Pred. No.: 17.9 Length: 1152
Score: 45.00 Matches: 9
Percent Similarity: 78.57% Conservative: 2
Best Local Similarity: 64.29% Mismatches: 3
Query Match: 50.00% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-15 (1-17) x US-09-938-842A-2486 (1-1152)
QY 4 LeuProLeuAspLysAspTyrTyrValValArgGluProGly 17
DB 952 CTTGACCTTGATGAGATGTTATGTTGACCAAGCCAGGT 993

RESULT 8
US-09-764-877-3727/C
; Sequence 3727, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3727
; LENGTH: 22786
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3727

Alignment Scores:
Pred. No.: 683 Length: 22786
Score: 45.00 Matches: 7
Percent Similarity: 84.62% Conservative: 4
Best Local Similarity: 53.85% Mismatches: 2
Query Match: 50.00% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x US-09-764-877-3727 (1-22786)
QY 4 LeuProLeuAspLysAspTyrTyrValValArgGluPro 16
DB 6948 ATCCCGCCGATGAGACTATATACATTGTTAGAAACCT 6910

RESULT 9
US-09-960-352-1663
; Sequence 1663, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nenping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NOCLETIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 1663
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 08-LIB188-026-Q1-E1-B7
```

US-09-960-352-1663

## Alignment Scores:

Pred. No.:	7.88	Length:	405
Score:	44.00	Matches:	7
Percent Similarity:	78.57%	Conservative:	4
Best Local Similarity:	50.00%	Mismatches:	3
Query Match:	48.89%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-15 (1-17) x US-09-960-352-1663 (1-405)

OY 4 LeuProLeuAspLysAspTyrTyrValValArgGluProGly 17

Db 229 CTGCCTGTCCACCAACACGTTACTCTCTGAAGACCCCGGT 270

## RESULT 10

US-09-960-352-11844/c  
Sequence 11844, Application US/09960352  
Patent No. US20020137139A1  
GENERAL INFORMATION:  
APPLICANT: Warren, Wesley C.  
APPLICANT: Tao, Nengbing  
APPLICANT: Byatt, John C.  
APPLICANT: Mathialagan, Nagappan  
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
FILE REFERENCE: 16511.006/37-21(10298)C  
CURRENT APPLICATION NUMBER: US/09/960.352  
CURRENT FILING DATE: 2001-09-24  
NUMBER OF SEQ ID NOS: 15112  
SEQ ID NO 11844  
LENGTH: 411  
TYPE: DNA  
ORGANISM: Bos taurus  
OTHER INFORMATION: Clone ID: 51-LIB188-010-01-E1-E4  
US-09-960-352-11844

## Alignment Scores:

Pred. No.:	8.02	Length:	411
Score:	44.00	Matches:	7
Percent Similarity:	78.57%	Conservative:	4
Best Local Similarity:	50.00%	Mismatches:	3
Query Match:	48.89%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-15 (1-17) x US-09-960-352-11844 (1-411)

OY 4 LeuProLeuAspLysAspTyrTyrValValArgGluProGly 17

Db 211 CTGCCTGTCCACCAACACGTTACTCTCTGAAGACCCCGGT 170

## RESULT 11

US-09-736-457-48  
Sequence 48, Application US/09736457  
Patent No. US20020168637A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Tonglong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Darriek  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fan, Liqun  
APPLICANT: Wang, Aijun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.478C15  
CURRENT APPLICATION NUMBER: US/09/736.457  
CURRENT FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 1864

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 48  
LENGTH: 430  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-736-457-48

Alignment Scores:	8.48	Length:	430
Pred. No.:	44.00	Matches:	11
Score:	68.42%	Conservative:	2
Best Local Similarity:	57.89%	Mismatches:	4
Query Match:	48.89%	Indels:	2
DB:	9	Gaps:	1

US-09-397-967-15 (1-17) x US-09-736-457-48 (1-430)

OY 1 AlalysLeuPro-----LeuAspLysAspTyrTyrValValArgGluProGly 17

Db 159 GCAAGCTTCCTCCAGTATTAGACCGATCTCTATTATTATTAAAGTGCT 215

## RESULT 12

US-09-902-941-48  
Sequence 48, Application US/09902941  
Patent No. US20020172952A1  
GENERAL INFORMATION:  
APPLICANT: Henderson, Robert A.  
APPLICANT: Wang, Tonglong  
APPLICANT: Watanabe, Yoshihiro  
APPLICANT: Johnson, Jeffrey C.  
APPLICANT: Retter, Marc W.  
APPLICANT: Marnerakis, Margarita  
APPLICANT: Carter, Darriek  
APPLICANT: Fanger, Gary R.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: McNabb, Andria  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.478C17  
CURRENT APPLICATION NUMBER: US/09/902.941  
CURRENT FILING DATE: 2001-07-10  
NUMBER OF SEQ ID NOS: 2002  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 48  
LENGTH: 430  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-902-941-48

## Alignment Scores:

Pred. No.:	8.48	Length:	430
Score:	44.00	Matches:	11
Percent Similarity:	68.42%	Conservative:	2
Best Local Similarity:	57.89%	Mismatches:	4
Query Match:	48.89%	Indels:	2
DB:	9	Gaps:	1

US-09-397-967-15 (1-17) x US-09-902-941-48 (1-430)

OY 1 AlalysLeuPro-----LeuAspLysAspTyrTyrValValArgGluProGly 17

Db 159 GCAAGCTTCCTCCAGTATTAGACCGATCTCTATTATTATTAAAGTGCT 215

## RESULT 13

US-09-849-626-48  
Sequence 48, Application US/09849626  
Publication No. US20020197659A1  
GENERAL INFORMATION:  
APPLICANT: Bangur, Chaitanya  
APPLICANT: Fanger, Gary  
APPLICANT: Wang, Aijun  
APPLICANT: Wang, Tonglong

```

; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-849-626-48

Alignment Scores:
Pred. No.: 8.48 Length: 430
Score: 44.00 Matches: 11
Percent Similarity: 68.42% Conservative: 2
Best Local Similarity: 57.89% Mismatches: 4
Query Match: 48.89% Indels: 2
DB: 9 Gaps: 1

US-09-397-967-15 (1-17) x US-09-849-626-48 (1-430)

QY 1 AlalysleuleuPro-----leuasplysasptyrtyrvalaIargluProgly 17
DB 159 GCAAGCTTCTCCCTCAGTATTAGACCGAGATCTCTATATTGTTAATGCTGT 215

RESULT 14
US-10-017-754-48
; Sequence 48, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-48

Alignment Scores:
Pred. No.: 8.48 Length: 430
Score: 44.00 Matches: 11
Percent Similarity: 68.42% Conservative: 2
Best Local Similarity: 57.89% Mismatches: 4
Query Match: 48.89% Indels: 2
DB: 9 Gaps: 1

US-09-397-967-15 (1-17) x US-10-017-754-48 (1-430)

QY 1 AlalysleuleuPro-----leuasplysasptyrtyrvalaIargluProgly 17
DB 159 GCAAGCTTCTCCCTCAGTATTAGACCGAGATCTCTATATTGTTAATGCTGT 215
```

```

RESULT 15
US-09-998-598-1085
; Sequence 1085, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stoik, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Cheneault, Ruth A.
; APPLICANT: Meagher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 1085
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-1085

Alignment Scores:
Pred. No.: 8.48 Length: 430
Score: 44.00 Matches: 11
Percent Similarity: 68.42% Conservative: 2
Best Local Similarity: 57.89% Mismatches: 4
Query Match: 48.89% Indels: 2
DB: 10 Gaps: 1

US-09-397-967-15 (1-17) x US-09-998-598-1085 (1-430)

QY 1 AlalysleuleuPro-----leuasplysasptyrtyrvalaIargluProgly 17
DB 159 GCAAGCTTCTCCCTCAGTATTAGACCGAGATCTCTATATTGTTAATGCTGT 215

Search completed: April 28, 2003, 18:56:43
Job time : 10.6765 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 28, 2003, 13:05:49 ; Search time 9.06362 Seconds

(without alignments)  
4223.916 Million cell updates/sec

Title: US-09-397-967-15

Perfect score: 90

Sequence: 1 AKLPLDKDYVYVREPG 17

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO-spool/US0937967/runat_26042003_182312_10702/app.query.fasta_1.1486
-DB=N_Geneseq_101002 -OFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -STRAP=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcr -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0937967.ecgn_1.1_418_etrunat_26042003_182312_10702 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSELOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database: N\_Geneseq\_101002:\*

```
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	100.0	165	16	AA087512	Tyrosine kinase pe
2	90	100.0	778	22	AA087512	Human Janus kinase
3	90	100.0	3607	17	AA087512	Protein tyrosine k
4	83	92.2	4016	17	AA087512	Mouse JAK3 protein
5	64	71.1	3435	19	AA087512	JAK2 protein encod
6	64	71.1	3473	13	AA087512	JAK2 encoding DNA
7	64	71.1	3495	24	AA087512	Murine JAK2 tyrosi
8	64	71.1	3629	16	AA087512	Murine JAK2 kinase
9	64	71.1	3629	21	AA087512	JAK2 polynucleotid
10	64	71.1	3629	22	AA087512	Murine JAK2 kinase
11	64	71.1	4482	20	AA087512	Murine (Janus) kin
12	64	71.1	4482	21	AA087512	Human JAK2 kinase
13	64	71.1	5117	24	AA087512	Human JAK2 kinase
14	51.5	57.2	1289	23	ABL21107	Human JAK2 (Janus
15	51.5	57.2	2743	23	ABL21107	Drosophila melanog
16	51.5	57.2	3419	23	ABL21107	Drosophila melanog
17	51.5	57.2	3530	23	ABL21107	Drosophila melanog
18	49	54.4	835	24	ABN85718	Human membrane-ass
19	48	53.3	32190	22	AA087512	Human cardiovascu
20	46	51.1	668	23	AA087512	Drosophila melanog
21	46	51.1	687	23	ABL24663	Human FRK1 gene.
22	46	51.1	1863	18	AA087512	Drosophila melanog
23	46	51.1	2065	23	AA087512	DNA encoding novel
24	46	51.1	2144	23	AA087512	DNA encoding novel
25	46	51.1	2747	23	ABL24662	Drosophila melanog
26	45	50.0	481	23	ABK42495	Genomic sequence #
27	45	50.0	861	24	ABN98821	Arabidopsis thalia
28	45	50.0	1152	21	AA087512	Arabidopsis thalia
29	45	50.0	1360	22	AA087512	Arabidopsis thalia
30	45	50.0	1375	21	AA087512	Arabidopsis thalia
31	45	50.0	1375	21	AA087512	H. vulgare Hcpl en
32	45	50.0	22786	21	AA087512	Arabidopsis thalia
33	45	50.0	22786	22	AA087512	Arabidopsis thalia
34	44	48.9	430	22	AA087512	Human muscleovale
35	44	48.9	430	22	AA087512	Human lung tumor
36	44	48.9	1143	24	ABK38041	CDNA encoding clon
37	44	48.9	1165	23	AA087512	Mouse G protein-co
38	44	48.9	1266	24	ABL50663	Rat r2A02 G protei
39	44	48.9	1387	21	AA087512	Lung cancer associ
40	44	48.9	1387	21	AA087512	Arabidopsis thalia
41	44	48.9	2013	21	AA087512	Arabidopsis thalia
42	44	48.9	2068	21	AA087512	Arabidopsis thalia
43	44	48.9	2070	21	AA087512	Arabidopsis thalia
44	44	48.9	3147	18	AA087512	Bovine lysosomal a
45	44	48.9	3147	18	AA087512	Bovine lysosomal a

## ALIGNMENTS

RESULT 1  
AA087512  
AA087512 standard; cDNA to mRNA; 165 BP.

22-NOV-1995 (first entry)

Tyrosine kinase peptide coding sequence.

Tyrosine kinase peptide: YF-7 cell; primer: polymerase chain reaction;  
PCR: cell differentiation; antisense drug; ds.

Homo sapiens.

JPO7059569-A.

07-MAR-1995.

```

XX PF 25-AUG-1993; 93JP-0210404.
XX PR 25-AUG-1993; 93JP-0210404.
XX PA (ASAH ) ASAH KASEI KOGYO KK.
XX DR WPI: 1995-135894/18.
XX DR P-PSDB: AAR71394.
XX PT New DNA sequence encoding a tyrosine kinase peptide - useful for
XX evaluation and control of cell differentiation.
XX PS Claim 1: Page 5: 5pp; Japanese.
XX CC This sequence encodes a tyrosine kinase peptide which is derived
XX from UT-7 cells. This sequence was isolated using the primer
XX CC sequences given in AA087510-11. The tyrosine kinase peptide may be
XX CC used for the evaluation and control of cell differentiation and is
XX CC also useful for the development of drugs and antisense drugs.
XX SQ Sequence 165 BP; 36 A; 56 C; 39 G; 34 T; 0 other:

Alignment Scores:
Pred. No.: 2,29e-08 Length: 165
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-397-967-15-(1-17) x AA087512 (1-165)
OY 1 AlalysleuProleuAspLysAspTyrTyrValAlaArgLupProGly 17
Db 52 GCTAAGCTGCTGCGCTTGACAAAGACTACTACTGCTGCGCGAGCCAGGC 102

RESULT 2
AAS10809
ID AAS10809 standard; cDNA; 778 BP.
XX AC AAS10809;
XX DT 24-OCT-2001 (first entry)
XX DE Human Janus kinase 3 (JAK3) partial cDNA sequence, 3-2 primer 1.
XX KW Human; Janus kinase 3; JAK3; ss; JAK/STAT inhibitor; 3-2 primer 1;
XX KW signal transducer and activator of transcription; osteoarthritis;
XX KW degenerative joint disease; rheumatoid arthritis; leprosy; asthma;
XX KW cancer; tumour; leukemia.
XX OS Homo sapiens.
XX PN WO200152892-A2.
XX PD 26-JUL-2001.
XX PF 22-JAN-2001; 2001WO-US02033.
XX PR 24-JAN-2000; 2000US-0177872.
XX PR 28-NOV-2000; 2000US-0723490.
XX PA (GENZ ) GENZYME CORP.
XX PI Vastios G;
XX DR WPI: 2001-465338/50.
XX PT Use of inhibitors of Janus, kinase/signal transducers and activators of
XX PT transcription for inhibiting onset and progression of degenerative
XX PT joint diseases or disorders such as osteoarthritis, rheumatoid
XX PT arthritis

```

```

XX PS Example 4; Fig 3; 55pp; English.
XX CC The sequence is a partial sequence, designated "clone 3-2 primer 1",
XX CC obtained from a chondrocyte cDNA encoding human Janus kinase
XX CC 3 (JAK3). The invention relates to the use of JAK/STAT (Janus
XX CC kinase/signal transducer and activator of transcription)
XX CC inhibitors other than debromylenaldisine (DBM) and
XX CC hylenaldisine (H) for inhibiting the progression or the likelihood of
XX CC developing diseases involving cartilage degradation, and for regulating
XX CC the expression of pro-inflammatory agents or cytokines in a chondrocyte,
XX CC and cartilage degrading enzymes in a cell. A JAK/STAT inhibitor of the
XX CC invention is useful for inhibiting progression or likelihood of
XX CC developing osteoarthritis or rheumatoid arthritis. The inhibitor is also
XX CC useful for treating other JAK/STAT-mediated diseases or disorders,
XX CC including T cell-mediated disorders, mast cell-mediated disorders,
XX CC type 2 (cytokine hypersensitivity) disorders, human T cell
XX CC myeloid diseases, T cell-mediated disorders include human T cell
XX CC leukemia/lymphoma virus (HTLV)-1, Sdzory's syndrome, c-abl
XX CC transformation, natural killer-like T cell lymphomas, (NK-like tumours)
XX CC and graft-vs-host disease; cytokine hypersensitivity disorders include
XX CC leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated
XX CC disorders include hay fever, asthma, hives and anaphylaxis; and
XX CC leukaemias and lymphomas include acute lymphocytic and lymphoblastic
XX CC leukaemias, B cell lymphomas and leukaemias of myeloid origin. DBM and H
XX CC are useful as therapeutic agents in cancers in which JAK3 plays a role
XX CC in the initiation or progression of tumorigenesis.
XX SQ Sequence 778 BP; 131 A; 264 C; 220 G; 163 T; 0 other:

Alignment Scores:
Pred. No.: 1.41e-07 Length: 778
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-397-967-15 (1-17) x AAS10809 (1-778)
OY 1 AlalysleuProleuAspLysAspTyrTyrValAlaArgLupProGly 17
Db 187 GCTAAGCTGCTGCGCTTGACAAAGACTACTACTGCTGCGCGAGCCAGGC 237

RESULT 3
AAT30862
ID AAT30862 standard; DNA; 3807 BP.
XX AC AAT30862;
XX DT 13-SEP-1996 (first entry)
XX DE Protein tyrosine kinase JAK3 cDNA.
XX KW JAK3; protein tyrosine kinase; cell proliferation; differentiation;
XX KW signal transduction; leukaemia; aplastic anaemia; myelodysplasia;
XX KW polythemia vera; thrombocytosis; gene therapy; diagnosis; ss.
XX OS Homo sapiens.
XX PN WO9618639-A1.
XX PD 20-JUN-1996.
XX PF 15-DEC-1995; 95WO-US16435.

Location/Qualifiers
Key 168..3359
FT CDS /*tag= a
FT polyA_signal 3692..3696
FT /*tag= b
FT polyA_signal 3793..3797
FT /*tag= c

```

```

XX 15-DEC-1994; 94US-0357598.
PR (0300) UNIV JOHNS HOPKINS SCHOOL MED.
PA CIVILN CT, Safford MG, Small D;
XX WPI: 1996-300568/30.
XX P-PSDB: AAR96037.
XX
XX Protein tyrosine kinase, JAK3, protein and nucleic acid - used in
PT the gene therapy of cellular proliferative diseases, e.g. leukaemia,
PT aplastic anaemia etc.
XX
XX Claim 4: Page 43-47; 97pp; English.
XX
XX A cDNA clone (AAT30862) codes for JAK3 (AAR96037), a new member of the
CC JAK family of non-receptor protein tyrosine kinases, that probably
CC plays a role in growth factor modulated differentiation,
CC proliferation and survival of haematopoietic stem/progenitor cells.
CC It was derived from CD34+ mRNA obtd. from normal human bone marrow
CC by PCR amplification using primers (see also AAT30863-64) based
CC on highly conserved motifs from protein tyrosine kinase catalytic
CC domains. The JAK3 coding sequence can be incorporated into a
CC vector and used for prodn. of recombinant JAK3. It can also be
CC used in gene therapy protocols for leukaemia, myelodysplasia,
CC polycythemia vera, thrombocytosis and aplastic anaemia, or to
CC stimulate haematopoietic cell proliferation. The gene was
CC localised to chromosome 19, band p12-13.1.
XX
XX Sequence 3807 BP; 716 A; 1219 C; 1099 G; 773 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 9,06e-07 Length: 3807
Score: 90.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
US-09-397-967-15 (1-17) x AAT30862 (1-3807)
QY 1 AAlaLysLeuLeuProLeuAspLysAspTyrTyrValValArgGluProGly 17
DB 3069 GCTAAAGCTGCTGCGCCCTGACAAAGACTACTAGCTGTCGCGAGCCAGCC 3119
RESULT 4
AAT11083
ID AAT11083 standard; cDNA: 4016 BP.
XX
XX AAT11083;
AC
XX
XX 08-APR-1996 (first entry)
DT
XX
XX Mouse JAK3 protein-tyrosine-kinase gene.
DE
XX
XX Mouse; Janus kinase; JAK3; protein-tyrosine-kinase; cytokine;
KW signal transduction; 32Dc13; reverse transcription; PCR;
KW polymerase chain reaction; primer; interleukin-3;
KW granulocyte-macrophage colony stimulating factor; cDNA probe;
KW granulocyte colony stimulating factor; Escherichia coli;
KW haematopoietic cell; differentiation; leukaemia; therapy; ss.
XX
XX Mus musculus.
OS
XX
XX Key Location/Qualifiers
FH CDS 7..3906
FT /*tag=a
FT /product= JAK3 protein-tyrosine-kinase
FT /note="EC-2.7.1.112"
XX
XX MO9601838-A1.
XX

```

```

PD 25-JAN-1996.
XX
XX 28-JUN-1995; 95MO-US08354.
XX
XX 08-JUL-1994; 94US-0272368.
XX
XX (UTEM) UNIV TEMPLE.
XX
XX Rane SG, Reddy PE;
PI WPI: 1996-097584/10.
XX P-PSDB: AAR8550.
XX
XX JAK3 protein tyrosine kinase and DNA encoding it - useful to induce
PT differentiation of haematopoietic cells to treat or control
PT leukaemia
XX
XX Claim 4: Page 26-28; 49pp; English.
XX
XX The sequence encodes a novel protein-tyrosine-kinase of the Janus
CC kinase (JAK) family (involved in cytokine signal transduction),
CC designated JAK3. The sequence has been isolated from mouse 32Dc13
CC cells growing in interleukin-3 or granulocyte-macrophage colony
CC stimulating factor by reverse transcription-polymerase chain
CC reaction amplification, using primers J3 and J4 (AAT11084-T11085),
CC based on conserved sequences in JAK JH2 and JH1 domains. A 950-bp
CC cDNA probe is isolated, which is then used to screen a granulocyte
CC colony stimulating factor-treated 32Dc13 cDNA library in phage
CC lambda-gt11 in Escherichia coli to isolate the full-length gene.
CC The JAK3 protein is useful for stimulating the differentiation of
CC haematopoietic cells in therapy or control of leukaemia.
XX
XX Sequence 4016 BP; 742 A; 1203 C; 1235 G; 836 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 2,05e-05 Length: 4016
Score: 83.00 Matches: 16
Percent Similarity: 94.12% Conservative: 0
Best Local Similarity: 94.12% Mismatches: 1
Query Match: 92.22% Indels: 0
DB: 17 Gaps: 0
US-09-397-967-15 (1-17) x AAT11083 (1-4016)
QY 1 AAlaLysLeuLeuProLeuAspLysAspTyrTyrValValArgGluProGly 17
DB 3544 GCTAAAGCTGCTGCGCCCTGCGAAGGACTACTAGCTGTCGCGAGCCCTGCG 3594
RESULT 5
AAV61801
ID AAV61801 standard; cDNA: 3435 BP.
XX
XX AAV61801;
AC
XX
XX 20-JAN-1999 (first entry)
DT
XX
XX JAK2 protein encoding cDNA.
DE
XX
XX JAK2 protein; cytoplasmic domain; beta subunit; screening; asthma;
KW interleukin; granulocyte macrophage-colony stimulating factor; GM-CSF;
KW IL-3; IL-5; human; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 10..3429
FT /*tag=a
FT /product= "JAK2 protein"
XX
XX MO9843087-A1.
XX
XX 01-OCT-1998.
XX

```

```

PF 23-MAR-1998; 98WO-US05387.
XX
XX 24-MAR-1997; 97US-0041511.
XX
XX (PHNA ) PHARMACIA & UPJOHN CO.
XX
XX Kaytes PS, Roberds SL;
XX
XX WPI: 1998-532151/45.
XX
XX P-PSDB: AAW76425.
XX
XX Screening for compounds useful for preventing or treating asthma -
XX by determining if compounds inhibit binding of the JAK2 protein to
XX e.g. IL-3, IL-5 or GM-CSF
XX
XX Disclosure: Fig 22A-C; 112pp; English.
XX
XX This cDNA encodes the full-length JAK2 protein. The N-terminal fragment
XX of JAK2 protein can be used in the method of the invention of screening
XX for compounds useful for treating or preventing asthma. The method
XX comprises contacting a molecule comprising at least the N-terminal 294
XX amino acid residues of the JAK2 protein, with another molecule comprising
XX at least 13 membrane-proximal cytoplasmic amino acids of interleukin
XX (IL)-3, IL-5 or granulocyte macrophage-colony stimulating factor (GM-CSF)
XX proteins in the presence of the candidate compound, and determining
XX whether the first and the second molecules form a complex. If the
XX compound inhibits complex formation, it can be used to treat asthma.
XX
XX Sequence 3435 BP; 1157 A; 598 C; 733 G; 947 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.0682 Length: 3435
XX Score: 64.00 Matches: 11
XX Percent Similarity: 87.50% Conservative: 3
XX Best Local Similarity: 68.75% Mismatches: 2
XX Query Match: 71.11% Indels: 0
XX DB: 19 Gaps: 0
XX
XX US-09-397-967-15 (1-17) x AAV61801 (1-3435)
XX
XX QY 2 LysleuleuProleuasplysAspTyrTyryValValArgIuProgly 17
XX |||||:||||| |||||:||||| |||||:|||||
XX Db 2998 AAGGCTTGCCACAGACAAAGATATCTATAAGTAAAGAACCTGGT 3045
XX
XX RESULT 6
XX AAQ25307
XX ID AAQ25307 standard; DNA; 3473 BP.
XX
XX AC AAQ25307;
XX
XX DT 18-FEB-1999 (first entry)
XX
XX DE JAK2 encoding DNA.
XX
XX DE JAK2 encoding DNA.
XX
XX KW Phosphorylation; JAK1; JAK2; protein tyrosine kinase; human;
XX catalytic domain; SH2 domain; growth factor receptor; PTK; murine; ss.
XX
XX OS Mus musculus.
XX
XX Key Location/Qualifiers
XX FT cds 1..2961
XX FT /*tag= a
XX FT /product= JAK2
XX
XX PN WO9210519-A.
XX
XX PD 25-JUN-1992.
XX
XX XX 26-NOV-1991; 91WO-US08889.
XX
XX PF 26-NOV-1991; 91WO-US08889.
XX
XX PR 28-NOV-1990; 90AU-0003594.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX

```

```

XX
XX Harpur A, Wilks AF, Ziemiecki A;
XX
XX WPI: 1992-234591/28.
XX
XX P-PSDB: AAR25141.
XX
XX Novel protein tyrosine kinase mol. - comprises multiple catalytic
XX domains but no SH2 domain and is for phosphorylation of proteins
XX
XX Claim 10; Fig 8; 50pp; English.
XX
XX This sequence encodes the murine protein tyrosine kinase JAK2 (from
XX Janus kinase). Northern analysis of JAK2 expression in a mouse
XX demonstrated the presence of two mRNA transcripts (4.8 and 4.4 kb).
XX The levels of these transcripts alter with respect to one another in
XX different tissues. The kidney, spleen and lung appear to express
XX predominantly the larger form, whereas ovary, placenta, skeletal muscle
XX and all murine cell lines analysed express both forms at equal levels.
XX The difference in sizes may be due to differential polyadenylation
XX sites. Both JAK2 and JAK1 are examples of a new subfamily or class
XX of protein tyrosine kinase. These can be used in the phosphorylation
XX of proteins, incorporation of labels and in the design of analogues,
XX antagonists and agonists of JAK's.
XX
XX Sequence 3473 BP; 1128 A; 677 C; 781 G; 887 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.0691 Length: 3473
XX Score: 64.00 Matches: 11
XX Percent Similarity: 87.50% Conservative: 3
XX Best Local Similarity: 68.75% Mismatches: 2
XX Query Match: 71.11% Indels: 0
XX DB: 13 Gaps: 0
XX
XX US-09-397-967-15 (1-17) x AAQ25307 (1-3473)
XX
XX QY 2 LysleuleuProleuasplysAspTyrTyryValValArgIuProgly 17
XX |||||:||||| |||||:||||| |||||:|||||
XX Db 2566 AAGGCTTGCCACAGACAAAGATATCTATAAGTAAAGAGCCGAGG 2613
XX
XX RESULT 7
XX AAD22680
XX ID AAD22680 standard; cDNA; 3495 BP.
XX
XX AC AAD22680;
XX
XX DT 26-FEB-2002 (first entry)
XX
XX DE Murine JAK2 tyrosine kinase cDNA.
XX
XX DE Murine JAK2 tyrosine kinase cDNA.
XX
XX KW Murine; Src homology 2-Bbeta; SH2-Bbeta; neuroprotective; gene therapy;
XX cell differentiation; nerve regeneration; angiogenesis; embryogenesis;
XX cytosolic; antisense therapy; drug screening; cellular expression;
XX immunological disease; neurological disease; apoptosis; diabetes; cancer;
XX arthritis; JAK2 tyrosine kinase; ss.
XX
XX OS Mus sp.
XX
XX Key Location/Qualifiers
XX FT CDS 1..2982
XX FT /*tag= a
XX FT /product= "murine JAK2 tyrosine kinase protein"
XX
XX PN US6312941-B1.
XX
XX PD 06-NOV-2001.
XX
XX XX 26-NOV-1997; 97US-0980080.
XX
XX PF 26-NOV-1997; 97US-0980080.
XX
XX PR 26-NOV-1997; 97US-0980080.
XX
XX (UNMT ) UNIV MICHIGAN.
XX

```





Query Match	71.11%	Indels:	0
DB:	16	Gaps:	0
US-09-397-967-15 (1-17) x AA085412 (1-3629)			
Qy	2 LysLeuLeuProLeuaspLysAspTyrTyrValAlaArgGluProGly 17		
	:           :           :		
Db	3088 AAGGCTTGCCGACAGACAAAGAACTACAAAGTAAGAGGCCAGGG 3135		
RESULT 9			
AA066244	ID AAC66244 standard; cDNA; 3629 BP.		
XX	AAC66244;		
AC	AAC66244;		
XX	19-FEB-2001 (first entry)		
DT	19-FEB-2001 (first entry)		
XX	Jak2 polynucleotide sequence.		
DE	Jak2 polynucleotide sequence.		
XX	Jak2; kinase; cytokine; cellular response; inhibition; jak2; ss;		
KW	cell proliferation; erythrocytosis.		
KM	cell proliferation; erythrocytosis.		
XX	Unidentified.		
OS	Unidentified.		
XX	US6136595-A.		
PN	US6136595-A.		
XX	24-OCT-2000.		
PD	24-OCT-2000.		
XX	18-JUN-1996; 96US-0665574.		
PF	18-JUN-1996; 96US-0665574.		
XX	29-JUL-1994; 94US-0282012.		
PR	29-JUL-1994; 94US-0282012.		
XX	29-JUL-1993; 93US-0097997.		
PR	29-JUL-1993; 93US-0097997.		
XX	09-SEP-1993; 93US-0118968.		
PA	(STUD-) ST JUDE CHILDREN'S RES HOSPITAL.		
XX	(STUD-) ST JUDE CHILDREN'S RES HOSPITAL.		
PI	Silvennoinen O, Wiltuhm BA, Ihle J;		
XX	Silvennoinen O, Wiltuhm BA, Ihle J;		
DR	WPI: 2000-686080/67.		
XX	P-PSDB: AAB35719.		
PT	New DNA encoding Jak3 kinase is useful as cytokine regulator for		
XX	treating cell proliferation		
PS	Example 1; Fig 1; 100pp; English.		
XX	Example 1; Fig 1; 100pp; English.		
CC	This invention relates to DNA encoding a murine Jak3 protein. The amino		
CC	acid sequence of the Jak3 protein is given in AAB35715. The jak family		
CC	of kinases are involved in the cellular response to the binding of		
CC	cytokines to their respective receptors. Jak3 kinase mediated activation		
CC	of some cytokines through their phosphorylation in response to		
CC	cytokine-receptor binding. Inhibiting the activity of Jak3 kinase (at the		
CC	nucleic acid level with antisense sequences or ribozymes, or at the		
CC	protein level with antibodies, kinase inhibitors etc.) is used to treat,		
CC	or diagnose, diseases caused by excessive secretion of certain cytokines		
CC	e.g. excessive cell proliferation such as erythrocytosis. Alternatively,		
CC	the Jak3 protein and polynucleotide can be used to treat conditions		
CC	associated with defective Jak3 activity. The DNA sequence can be used to		
CC	produce recombinant Jak3 and thus used to raise antibodies useful as		
CC	specific inhibitors or to detect or isolate Jak3 without interfering		
CC	with its enzymatic activity. The present sequence represents cDNA		
CC	encoding a Jak2 protein used in the isolation and characterisation of the		
XX	Jak3 protein of the invention.		
XX	Jak3 protein of the invention.		
SO	Sequence 3629 BP; 1142 A; 732 C; 843 G; 912 T; 0 other;		
Alignment Scores:	Alignment Scores:		
Pred. NO.:	0.0728	Length:	3629
Score:	64.00	Matches:	11
Percent Similarity:	87.50%	Conservative:	2
Best Local Similarity:	68.75%	Mismatches:	3
Query Match:	71.11%	Indels:	0
DB:	21	Gaps:	0

US-09-397-967-15 (1-17)	x AAC66244 (1-3629)
QY	2 LysleuLeuProLeuAspLysAspTyrTyrValValarglsuProgly 17
DB	3088 AAGGCTTGCCGACAGACAAAGATACTACAAAGTAAGGACCGAGG 3135
RESULT 10	
AD03607	
ID	AD03607 standard; cDNA; 3629 BP.
XX	
AC	AD03607;
XX	
DT	19-JUN-2001 (first entry)
XX	
DE	Murine (Janus kinase 2) Jak2 cDNA.
XX	
KW	Mouse; Janus kinase 2; Jak2; therapy; cytokine; tyrosine kinase;
KW	antiproliferative; cyostatic; cell proliferative disorder;
KW	cellular response; erythrocytosis; ss.
XX	
XX	Mus sp.
Key	Location/Qualifiers
FH	1..93
FH	5'UTR
FT	/*tag= a
FT	94..3483
FT	/*tag= b
FT	/product= "Murine (Janus kinase 2) Jak2 protein"
FT	523
FT	/*tag= c
FT	/note= "end of published partial Jak2 cDNA
FT	sequence (Harpur et al.)"
FT	replace (551..553, CCC)
FT	/*tag= d
FT	replace (1089, T)
FT	/*tag= e
FT	replace (1103, C)
FT	/*tag= f
FT	replace (1111, G)
FT	/*tag= g
FT	replace (1119, G)
FT	/*tag= h
FT	replace (1122, C)
FT	/*tag= i
FT	replace (1128, C)
FT	/*tag= j
FT	replace (1131, G)
FT	/*tag= k
FT	replace (1134, G)
FT	/*tag= l
FT	replace (1137, C)
FT	/*tag= m
FT	replace (1140, G)
FT	/*tag= n
FT	replace (1143, G)
FT	/*tag= o
FT	replace (1146, C)
FT	/*tag= p
FT	replace (1188, T)
FT	/*tag= q
FT	replace (1194, G)
FT	/*tag= r
FT	replace (1230, G)
FT	/*tag= s
FT	replace (1245, T)
FT	/*tag= t
FT	replace (1260, T)
FT	/*tag= u
FT	replace (1266, C)
FT	/*tag= v
FT	replace (1272, T)
FT	/*tag= w
FT	conflict

```

FT conflict replace (1275, Q)
FT /*tag- x
FT conflict replace (1293, T)
FT /*tag- y
FT conflict replace (1305, T)
FT /*tag- z
FT conflict replace (1323, C)
FT /*tag- aa
FT conflict replace (1341, A)
FT /*tag- ab
FT conflict replace (1344, A)
FT /*tag- ac
FT conflict replace (1359, G)
FT /*tag- ad
FT conflict replace (1365, Q)
FT /*tag- ae
FT conflict replace (1368, T)
FT /*tag- af
FT conflict replace (1374, T)
FT /*tag- ag
FT conflict replace (1401, C)
FT /*tag- ah
FT conflict replace (1413, C)
FT /*tag- ai
FT conflict replace (1431, T)
FT /*tag- aj
FT conflict replace (1453, Q)
FT /*tag- ak
FT conflict replace (1476, G)
FT /*tag- al
FT conflict replace (1488, T)
FT /*tag- am
FT conflict replace (1511..1512, GT)
FT /*tag- an
FT conflict replace (1578, C)
FT /*tag- ao
FT conflict replace (1590, T)
FT /*tag- ap
FT conflict replace (1593, T)
FT /*tag- aq
FT conflict replace (1602, G)
FT /*tag- ar
FT conflict replace (1623, T)
FT /*tag- as
FT conflict replace (1642, G)
FT /*tag- at
FT conflict replace (1657, C)
FT /*tag- au
FT conflict replace (1728, G)
FT /*tag- av
FT conflict replace (1743, C)
FT /*tag- aw
FT conflict replace (1755, C)
FT /*tag- ax
FT conflict replace (1770, A)
FT /*tag- ay
FT conflict replace (1809, G)
FT /*tag- az
FT conflict replace (1816, G)
FT /*tag- ba
FT conflict replace (1821, C)
FT /*tag- bb
FT conflict replace (1857, A)
FT /*tag- bc
FT conflict replace (1878, T)
FT /*tag- bd
FT conflict replace (1935, A)
FT /*tag- be
FT conflict replace (1938, A)
FT /*tag- bf
FT conflict replace (1963, T)
FT /*tag- bg
FT conflict replace (1974, G)

```

```

FT conflict /*tag- bh
FT conflict replace (2025, T)
FT conflict /*tag- bi
FT conflict replace (2055, G)
FT conflict /*tag- bj
FT conflict replace (2079, C)
FT conflict /*tag- bk
FT conflict replace (2082, C)
FT /*tag- bl
FT misc_feature 2226
FT /*tag- bm
FT /*note= "Nucleotides corresponding to 7 amino acid insert"
FT conflict replace (2253, A)
FT /*tag- bn
FT conflict replace (2259, G)
FT /*tag- bo
FT conflict replace (2283, A)
FT /*tag- bp
FT conflict replace (2285, C)
FT /*tag- bq
FT conflict replace (2433, G)
FT /*tag- br
FT conflict replace (3453, C)
FT /*tag- bs
FT 3'UTR 3484..3629
FT /*tag- bt
FT conflict replace (3579, C)
FT /*tag- bu
FT conflict replace (3595, A)
FT /*tag- bv
FT conflict replace (3599, A)
FT /*tag- bw
FT conflict replace (3620, A)
FT /*tag- bx

```

```

PN US6210654-B1.
PD 03-APR-2001.
PF 08-OCT-1997; 97US-0946994.
PR 18-JUN-1996; 96US-0665574.
PR 29-JUL-1993; 93US-0097997.
XX (SUD-) ST JUDE CHILDREN'S HOSPITAL.
PA Ihle J, Witthuhn BA, Quelle FW, Silvennoinen O;
XX WPI: 2001-265367/27.
XX P-PSDB: AAE00352.
XX
XX Modulating a biological response mediated by Jak kinase 2 activation to
XX a cytokine, useful for treating excessive proliferation of eukaryotic
XX cells, comprises inhibiting or enhancing tyrosine kinase activity of
XX Jak kinase in the cell.
XX
XX Example 1: Fig 1; 100pp; English.
XX
XX The present sequence is a cDNA encoding murine (Janus kinase 2) Jak2
XX tyrosine kinase. Jak2 sequence has a 600 amino acid long N-terminus that
XX lacks obvious SH2 (src homology 2) and SH3 domains. Following this is a
XX kinase related domain (domain 2) and a carboxyl kinase domain (domain
XX 1). Jak kinases mediate cytokine activity through their tyrosine

```

## Alignment Scores:

```

Pred. No.: 0.0728 Length: 3629
Score: 64.00 Matches: 11
Percent Similarity: 87.50% Conservative: 3
Best Local Similarity: 68.75% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 22 Gaps: 0

```

US-09-397-967-15 (1-17) x AAD03607 (1-3629)

OY 2 LysLeuLeuProLeuAspTyrTyrValValArgGluProGly 17  
 |||:::||||| |||:::||||| |||:::|||||  
 DB 3088 AAGGCTTGCCGACAGCAAGAAATACAAAGTAAAGAGCCAGG 3135

RESULT 11  
 AAX80971  
 ID AAX80971 standard; cDNA: 4482 BP.  
 AC AAX80971;  
 XX  
 XX 03-SEP-1999 (first entry)  
 DE Human JAK2 kinase encoding cDNA.  
 XX  
 XX JAK2 kinase; arteriosclerosis; asthma; bronchitis; emphysema; psoriasis;  
 KW inflammatory bowel disease; inflammation; osteoarthritis; oncogenesis;  
 KW rheumatoid arthritis; septic shock; systemic lupus erythematosus;  
 KW leukaemia; human; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX US5914393-A.  
 PN  
 XX 22-JUN-1999.  
 PD  
 XX 05-DEC-1995; 95US-0567508.  
 PF  
 XX 05-DEC-1995; 95US-0567508.  
 PR  
 XX (INCY-) INCYTE PHARM INC.  
 PA  
 XX Coleman R, Stuart SG;  
 PI  
 XX WPI: 1999-384188/32.  
 DR P-PSDB: AAY21698.  
 DR  
 XX Polypeptides and amino acids useful for modulating human jak2 kinase  
 PT activity  
 PT  
 XX  
 XX Claim 2; Columns 23-28; 37pp; English.  
 PS  
 XX This cDNA encodes a human JAK2 kinase polypeptide. Host cells transformed  
 CC with recombinant JAK2 kinase nucleic acid are used for the recombinant  
 CC production of the protein. Purified JAK2 may be used to produce  
 CC antibodies or identify antagonists or inhibitors of JAK2. JAK2, anti-JAK2  
 CC antibodies and JAK2 antagonists or inhibitors may be used to treat,  
 CC prevent or diagnose conditions associated with altered or uncontrolled  
 CC JAK2 expression, e.g. arteriosclerosis, asthma, bronchitis, emphysema,  
 CC inflammatory bowel disease, inflammation, leukemia, oncogenesis,  
 CC osteoarthritis, psoriasis, rheumatoid arthritis, septic shock and  
 CC systemic lupus erythematosus.  
 CC  
 CC  
 SQ Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;

Alignment Scores:  
 Pred. No.: 0.0932 Length: 4482  
 Score: 64.00 Matches: 11  
 Percent Similarity: 87.50% Conservative: 3  
 Best Local Similarity: 68.75% Mismatches: 2  
 Query Match: 71.11% Indels: 0  
 DB: 20 Gaps: 0

US-09-397-967-15 (1-17) x AAX80971 (1-4482)

OY 2 LysLeuLeuProLeuAspTyrTyrValValArgGluProGly 17  
 |||:::||||| |||:::||||| |||:::|||||  
 DB 3334 AAGGCTTGCCGACAGCAAGAAATACAAAGTAAAGAGCCAGG 3381

RESULT 12  
 AAZ58947  
 ID AAZ58947 standard; cDNA: 4482 BP.

AC AAZ58947;  
 XX  
 XX 03-MAY-2000 (first entry)  
 DE Human JAK2 kinase (HJAK2) encoding cDNA.  
 XX  
 XX Janus family nonreceptor protein-tyrosine kinase-2; JAK2; HJAK2; human;  
 KW signal transduction; arteriosclerosis; asthma; bronchitis; emphysema;  
 KW inflammatory bowel disease; leukemia; oncogenesis; osteoarthritis;  
 KW psoriasis; rheumatoid arthritis; systemic lupus erythematosus;  
 KW cytostatic; osteopathic; dermatological; antibacterial; septic shock;  
 KW immunosuppressive; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX US6019966-A.  
 PN  
 XX 01-FEB-2000.  
 PD  
 XX 19-NOV-1998; 98US-0196480.  
 PF  
 XX 05-DEC-1995; 95US-0567508.  
 PR  
 XX (INCY-) INCYTE PHARM INC.  
 PA  
 XX Stuart SG, Coleman R;  
 PI  
 XX WPI: 2000-146859/13.  
 DR P-PSDB: AAY77552.  
 DR  
 XX Human Janus family nonreceptor protein-tyrosine kinases useful as  
 PT diagnostic reagents and for preventing, diagnosing and treating  
 PT diseases such as arteriosclerosis, asthma and leukemia -  
 PT  
 XX Example 5; Fig 1A-F; 33pp; English.  
 PS  
 XX This cDNA encodes a human Janus family nonreceptor protein-tyrosine  
 CC kinase-2 (JAK2) polypeptide (HJAK2). The JAK2 polypeptides may be used  
 CC as diagnostic reagents as they react with a range of target proteins  
 CC including growth hormone, prolactin, erythropoietin and cytokine  
 CC receptors. They may also be used for the production of antibodies  
 CC specific for JAK2, which may be used to inhibit its activity and prevent  
 CC or treat disorders associated with over expression of JAK2. Conversely,  
 CC the JAK2 polypeptide may be administered to supplement the patients own  
 CC production and counter mutations that may lead to the expression of an  
 CC inactive enzyme. The protein may also be used to screen candidate  
 CC reagents for modulators of JAK2 function. The antagonists and antibodies  
 CC bind to the JAK2 protein and prevent the transfer of high energy  
 CC phosphate molecules, therefore blocking signal transduction. Disorders  
 CC that may be treated by administration of JAK2 polypeptides, anti-JAK2  
 CC antibodies and the agonists and/or antagonists, include arteriosclerosis,  
 CC asthma, bronchitis, emphysema, inflammatory bowel disease, leukemia,  
 CC osteoarthritis, psoriasis, inflammation, leukemia, oncogenesis,  
 CC shock and systemic lupus erythematosus.  
 CC  
 CC  
 SQ Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;

Alignment Scores:  
 Pred. No.: 0.0932 Length: 4482  
 Score: 64.00 Matches: 11  
 Percent Similarity: 87.50% Conservative: 3  
 Best Local Similarity: 68.75% Mismatches: 2  
 Query Match: 71.11% Indels: 0  
 DB: 21 Gaps: 0

US-09-397-967-15 (1-17) x AAZ58947 (1-4482)

OY 2 LysLeuLeuProLeuAspTyrTyrValValArgGluProGly 17  
 |||:::||||| |||:::||||| |||:::|||||  
 DB 3334 AAGGCTTGCCGACAGCAAGAAATACAAAGTAAAGAGCCAGG 3381

RESULT 13  
 AAD24311

ID AAD24311 standard; DNA; 5117 BP.  
 AC AAD24311;  
 XX  
 DT 07-MAR-2002 (first entry)  
 XX  
 XX Human Jak2 (Janus kinase) DNA.  
 DE  
 XX  
 KW Human; SOCS: suppressor of cytokine signaling; autoimmune disorder; Jak;  
 KW Janus kinase; signal transducer and activator of transcription; Stat;  
 KW transplant acceptance; systemic lupus erythematosus; glomerulonephritis;  
 KW rheumatoid arthritis; Wegener's granulomatosis; chronic active hepatitis;  
 KW atopic dermatitis; multiple sclerosis; myasthenia gravis; vasculitis;  
 KW Crohn's disease; hemolytic anaemia; nephrotic syndrome; dermatological;  
 KW diabetes mellitus; thyroiditis; inflammatory bowel disease; nephrotropic;  
 KW immunosuppressive; antineoplastic; antirheumatic; antiarthritic;  
 KW hepatotropic; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 478..3876  
 FT /tag= a  
 FT /product= "Human Jak2 protein"  
 PN WO200179555-A2.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 13-APR-2001; 2001WO-US12131.  
 XX  
 PR 14-APR-2000; 2000US-0549654.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Hancock MW, Ozkaymak E;  
 XX  
 DR WPI: 2002-034368/04.  
 DR P-PSDB: MAE15179.  
 XX  
 XX  
 PT Monitoring transplant acceptance or autoimmune disease, useful e.g. for  
 PT assessing therapy, comprises measuring levels of Stat or their  
 PT inhibitors  
 XX  
 PS Example: Fig 8; 218pp: English.  
 XX  
 CC The present invention relates to a method for monitoring acceptance of a  
 CC transplant or an autoimmune disease in a mammal. The method comprising  
 CC determining the amount of at least one of Stat4 (signal transducer and  
 CC activator of transcription), Stat6, SOCS1 (suppressor of cytokine  
 CC signaling) or SOCS3 mRNA or protein in a sample of the transplant, taken  
 CC from the host or an affected tissue sample. Stats are activated by  
 CC receptor-associated Janus kinases (Jaks) which include Jak1, Jak2, Tyk2,  
 CC Jak3. The method is used to determine whether acceptance of a transplant  
 CC has been induced or to determine if autoimmune disorders (systemic  
 CC lupus erythematosus, glomerulonephritis, rheumatoid arthritis, Wegener's  
 CC granulomatosis, chronic active hepatitis, atopic dermatitis, multiple  
 CC sclerosis, myasthenia gravis, hemolytic anaemia, nephrotic syndrome,  
 CC thyroiditis, diabetes mellitus, Crohn's disease, inflammatory bowel  
 CC disease and vasculitis) are being treated successfully and may be used  
 CC to adjust treatment regimes. The present sequence is human Jak2 DNA.  
 XX  
 SO Sequence 5117 BP; 1623 A; 927 C; 1103 G; 1464 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.109 Length: 5117  
 Score: 64.00 Matches: 11  
 Percent Similarity: 87.50% Conservative: 3  
 Best Local Similarity: 68.75% Mismatches: 2  
 Query Match: 71.11% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-397-967-15 (1-17) x AAD24311 (1-5117)

OY 2 LysLeuLeuProLeuAspLysAspTyrTyrValValArgLupProGly 17  
 DB 3472 AAGCTTCCACACAGCAAGAAAGTAATCTAATAAGTAAGAAAGCTGCT 3519  
 RESULT 14  
 ABL21107  
 ID ABL21107 standard; DNA; 1289 BP.  
 XX  
 AC ABL21107;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 14794.  
 DE XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 KW  
 KW Drosophila melanogaster.  
 OS  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI: 2001-656860/75.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 XX  
 PS Claim 1; SEQ ID NO 14794; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABBS5737-ABBS72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
 XX  
 SO Sequence 1289 BP; 274 A; 378 C; 304 G; 333 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 5.07 Length: 1289  
 Score: 51.50 Matches: 11  
 Percent Similarity: 77.78% Conservative: 3  
 Best Local Similarity: 61.11% Mismatches: 1  
 Query Match: 57.22% Indels: 3  
 DB: 23 Gaps: 1  
 US-09-397-967-15 (1-17) x ABL21107 (1-1289)  
 OY 3 LeuLeuProLeuAspLysAspTyrTyrValValArgLupProGly 17  
 DB 446 CTTCCACACACGACAGAGATATTACGTTACGTCGATCTGCTGCCGCGC 499  
 RESULT 15  
 ABL25944  
 ID ABL25944 standard; DNA; 2743 BP.  
 XX

AC ABL25944;

DT 26-MAR-2002 (first entry)

XX DE	Drosophila melanogaster genomic polynucleotide seq	ID NO 29305.

XX  
KW  
Drosophila; developmental biology; cell signalling; insecticide

pharmaceutical; gene; ds-

*Drosophila melanogaster*.

AA  
PN WO200171042-A2.

XX 27-SEP-2001.  
PD

XX  
PF 23-MAR-2001; 2001WO-US09231.

XX  
PR 23-MAR-2000; 2000US-191637P.

11-JUL-2000; 2000US-0614150.

AA (PEKE) PE CORP NY.  
PA  
NY

XX  
PI      Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.  
DR

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT isolated nucleic acid detection reagent for detecting 1000 or more

genes from *Drosophila* and for elucidating cell signalling and cell-cell interactions -

Claim 1: SEO ID NO 29305; 21pp + Sequence Listing; English.

XX a) mutation relates to an isolated nucleic acid detection reagent

CC The invention relates to an apparatus and a method for  
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
CC directed to a method and in elucidating cell signalling and

cell-cell interactions in higher eukaryotes for the development of useful in developmental biology and in understanding the mechanisms of drug action. The invention of cell-cell interactions in higher eukaryotes for the development of useful in developmental biology and in understanding the mechanisms of drug action.

CC insecticides, therapeutics and pharmaceutical drugs. The expressed DNA  
CC discloses genomic DNA sequences (AB16176-AB13051), expressed DNA  
CC discloses genomic DNA sequences (AB16176-AB13051), expressed DNA

sequences (ABL01840-ABL16175) and the encoded proteins (ABR57737-ABR72072).

The sequence data for this patent did not form part of the printed publication, but was obtained in electronic format directly from

CC Specification, see [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

XX	Sequence	2743 BP;	583 A;	768 C;	739 G;	653 T;	0 other;
sq							

Alignment Scores: 2743

Pred. No.:	12.3	Length:	27.5
Score:	51.50	Matches:	11
		Iterations:	3

Percent Similarity:	77.78%	Conservative:	3
Post Local Similarity:	61.1%	Mismatches:	1

Best local alignment:	57.228	Indels:	3
Query Match:	23	Gaps:	1

DB: 22  
APR 25 1964 (1-2743)

US-09-397-96/-15 (I-1) X ABLE2544 (1-1-15)

qy ||||| 3 LeuLeuProLeuasplysaspyrilyvalvalnigsl  
 : : : : :  
 ||||| 3638

Db 2573 CTTCTCCACTCGTACAGAGTATTACGTTATCCGTGAGTATCTGCTGCCCC 262

Search completed: April 28, 2003, 14:00:39  
Job time : 13.0636 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 28, 2003, 13:46:54 ; Search time 53.2545 Seconds

(without alignments)  
5169.955 Million cell updates/sec

Title: US-09-397-967-15

Perfect score: 90

Sequence: 1 AKLLPLDKDYVVRPG 17

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV-xlp  
-MODE=frame+pn.model -DEV-xlp  
-O/cgn2.1/USPICO/US09397967/runat.26042003.182313.10728/app.query.fasta.1.1486  
-DB=EST -OPMT=fastp -SUPER=first -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTEPT=pcr -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967@cgn.1.1.3008 -runat.26042003.182313.10728 -MCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

EST:

1: em\_estba:\*  
2: em\_estlum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estcp1:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_dln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fut:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	90	100.0	462	10	AA008832
2	90	100.0	544	12	BF705690
3	90	100.0	562	12	BF193178
4	86	95.6	417	13	BF540725
5	86	95.6	481	17	AO192794
6	86	95.6	543	12	BE753801
7	83	92.2	453	9	AA023709
8	83	92.2	478	9	AA023670
9	83	92.2	512	9	AA510093
10	83	92.2	593	9	AA881653
11	83	92.2	698	14	BM935182
12	83	92.2	720	12	BG873355
13	83	92.2	907	12	BG870513
14	80	88.9	592	9	A1981562
15	80	88.9	769	9	AJ442008
16	74	82.2	590	14	W48204
17	69	76.7	989	17	CNS042XQ
18	69	76.7	1001	17	CNS038MW
19	67	74.4	449	10	BB711016
20	66	73.3	867	9	AA755769
21	64	71.1	398	9	AA453345
22	64	71.1	532	12	BG364706
23	64	71.1	562	10	AW231037
24	64	71.1	753	12	BG870416
25	64	71.1	767	13	BT734179
26	64	71.1	920	14	BO898824
27	64	71.1	945	14	BO732700
28	58	64.4	258	12	BF367544
29	57	63.3	610	9	AU091091
30	57	63.3	612	9	AA521703
31	57	63.3	694	10	AA699477
32	51.5	57.2	687	9	AA539009
33	51.5	57.2	709	13	BT588497
34	51	56.7	141	9	AA014812
35	51	56.7	308	9	AA174246
36	51	56.7	434	17	AQ177309
37	51	56.7	766	13	BT153342
38	49	54.4	379	10	AV540081
39	49	54.4	437	17	AO009280
40	49	54.4	551	12	BT194671
41	49	54.4	569	17	AQ414068
42	49	54.4	651	10	BB620498
43	48.5	53.9	789	10	AW154977
44	48.5	53.9	851	17	CNS03HSP
45	48	53.3	319	9	AA358653

## ALIGNMENTS

RESULT 1  
AA008832  
LOCUS  
DEFINITION  
IMAGE:3063389 5', mRNA sequence.  
ACCESSION  
AA008832  
VERSION  
AA008832.1 GI:6927889  
KEYWORDS  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 (bases 1 to 462)  
NIH-MGC http://mgi.nci.nih.gov/.  
AUTHORS  
National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE  
462 bp mRNA linear EST 16-FEB-2000  
UT-HF-BM0-idx-h-03-0-UT-rl NIH-MGC-38 Homo sapiens cDNA clone

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: rgs@pshs-remail.nih.gov  
Eco RI site shown at the beginning of the sequence.  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: M.B. Soares Lab  
CDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E.E. Consortium/ILM at:  
www.bio.liml.gov/bdrp/limag/image.html  
Seq primer: M13 Forward.

FEATURES  
source

Location/Qualifiers  
1..462

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="3063389"  
/clone\_lib="NIH\_MGC\_38"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/cell\_line="MGC85"  
/lab\_host="DH10B (LTR)"  
/note="Vector: pTR73-Pac; Site\_1: NotI; Site\_2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA  
(2.5-3.5kb). Directionally cloned. Cells provided by Louis  
M. Staudt, Ph.D. Library preparation by Maria de Fatima  
Bonaldi, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT  
ORIGIN

86 a 156 c 131 g 89 t

## Alignment Scores:

Pred. No.: 2.24e-06 Length: 462  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-15 (1-17) x AM408832 (1-462)

Oy 1 AlalysleuProleuaspIyryValaValaArgluProgly 17

Db 349 GCTAGCTGCTCCGCTGACAAAGACTACTAGTGTGCGGACCAAGGC 399

## RESULT 2

LOCUS BF705690 544 bp mRNA linear EST 29-DEC-2000  
DEFINITION 243606 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
ACCESSION BF705690  
VERSION BF705690.1 GI:11997351  
KEYWORDS EST.

SOURCE  
plg.  
Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE  
AUTHORS  
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,  
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.  
and Keeler, J.W.

TITLE  
Design and use of two pooled tissue normalized cDNA libraries for  
EST discovery in swine  
Unpublished (2000)  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390

JOURNAL  
COMMENT

Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt-trimmed with phred  
v0.980904.e. Vector identified by cross-match with the -minscore 18  
and -minmatch 12 options.  
PCR primers  
FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTTCCAGTCACGACG  
Plate: 70 row: D column: 5  
Seq primer: ATTTAGGTGACACTATAG.

FEATURES  
source

Location/Qualifiers  
1..544

/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_lib="MARC 2P1G"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
Library made from pooled tissue from testis, ovary,  
endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT  
ORIGIN

96 a 181 c 156 g 111 t

## Alignment Scores:

Pred. No.: 2.82e-06 Length: 544  
Score: 90.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967-15 (1-17) x BF705690 (1-544)

Oy 1 AlalysleuProleuaspIyryValaValaArgluProgly 17

Db 154 GCGAGCTGCTCCGCTGACAAAGACTACTATGTGTGCGGACCGCGGC 204

## RESULT 3

BF193178 BF193178 562 bp mRNA linear EST 02-NOV-2000  
DEFINITION 244670 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
ACCESSION BF193178  
VERSION BF193178.1 GI:11076547  
KEYWORDS EST.

SOURCE  
plg.  
Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE  
AUTHORS  
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,  
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.  
and Keeler, J.W.

TITLE  
Design and use of two pooled tissue normalized cDNA libraries for  
EST discovery in swine  
Unpublished (2000)  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390

Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt-trimmed with phred  
v0.980904.e. Vector identified by cross-match with the -minscore 18  
and -minmatch 12 options.  
PCR primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTTCCAGTCACGACG  
Plate: 72 row: F column: 7  
Seq primer: ATTTAGGTGACACTATAG.

Location/Qualifiers  
1..562

FEATURES  
source

/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_lib="MARC 2P1G"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
Library made from pooled tissue from testis, ovary,  
endometrium, hypothalamus, pituitary, and placenta."

## BASE COUNT

97 a 188 c 167 g 110 t



ORIGIN

Alignment Scores:

Pred. No.:	2,95e-06	Length:	562
Score:	90.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-09-397-967-15 (1-17) x BF193178 (1-562)

OY 1 AlalysleuProleuAsplysAspyrTyValValArgLupProgly 17

DB 165 GCGAAGCTGCTGCCCTGACAAAGACTACTATGTGTGCGCGAGCCGCCG 215

RESULT 4

LOCUS B1540725 417 bp mRNA linear EST 30-AUG-2001

DEFINITION 454105 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION B1540725

VERSION B1540725.1 GI:15381837

KEYWORDS EST.

SOURCE

ORGANISM Bos taurus

COV.

REFERENCE

AUTHORS 1 (bases 1 to 417)

Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pettea,G., Holt,I., Karameycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

MEDLINE

COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.

PCR primers

FORWARD: AGGAACAGCATGACCAT

BACKWARD: GTTTCCAGTCACGACG

Plate: 118 row: P column: 4

Seq primer: ATTAGGTGACACTATAG.

FEATURES

source

1..417

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_lib="MARC 1BOV"

/tissue\_type="pooled"

/lab\_host="DH10B"

/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI; Library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."

BASE COUNT 74 a 138 c 121 g 83 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	1.04e-05	Length:	417
Score:	86.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	94.12%	Mismatches:	1
Query Match:	95.56%	Indels:	0
DB:	13	Gaps:	0

US-09-397-967-15 (1-17) x B1540725 (1-417)

OY 1 AlalysleuProleuAsplysAspyrTyValValArgLupProgly 17

DB 85 GCGAAGCTGCTGCCCTGACAAAGACTACTATGTGTGCGCGAGCCAGCC 135

RESULT 5

LOCUS A0192794/c 481 bp DNA linear GSS 04-NOV-1998

DEFINITION HS-2251\_B2\_F07\_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-2251 Col-14 Row-L, DNA sequence.

ACCESSION A0192794

VERSION A0192794.1 GI:3591416

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE

AUTHORS 1 (bases 1 to 481)

Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

MEDLINE

COMMENT Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 2251 row: L column: 14

Class: BAC ends

High quality sequence stop: 481.

FEATURES

source

1..481

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="Plate-2251 Col-14 Row-L"

/sex="male"

/note="Origin: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 101 a 115 c 192 g 72 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	1.27e-05	Length:	481
Score:	86.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	95.56%	Indels:	0
DB:	17	Gaps:	0

US-09-397-967-15 (1-17) x A0192794 (1-481)

OY 2 LysleuProleuAsplysAspyrTyValValArgLupProgly 17

DB 360 AAGCTGCTGCCCTGACAAAGACTACTATGTGTGCGCGAGCCAGCC 313

RESULT 6

LOCUS BE753801 543 bp mRNA linear EST 25-APR-2001

DEFINITION 206963 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BE753801

VERSION BE753801.1 GI:10167793

KEYWORDS EST.

SOURCE cow.

ORGANISM Bos taurus

REFERENCE

AUTHORS 1 (bases 1 to 543)

Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

MEDLINE

COMMENT Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 2251 row: L column: 14

Class: BAC ends

High quality sequence stop: 481.

FEATURES

source

1..543

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_lib="MARC 2BOV"

/tissue\_type="pooled"

/lab\_host="DH10B"

/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI; Library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."

BASE COUNT 74 a 138 c 121 g 83 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	1.04e-05	Length:	417
Score:	86.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	94.12%	Mismatches:	1
Query Match:	95.56%	Indels:	0
DB:	13	Gaps:	0

REFERENCE  
AUTHORS  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 543)  
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
Casas,E., Wray,D.E., White,J., Cho,U., Fahrenkrug,S.C., Bennett,  
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,  
Pettee,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and  
Keele,J.W.  
Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)  
21180013  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smithemall.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.  
PCR primers  
FORWARD: AGGAACAGCATGACCAT  
BACKWARD: GTTTCACGACGACGACG  
Plate: 49 row: M column: 16  
Seq primer: ATTAGGTGACACTATAG.  
Location/Qualifiers  
1. 543  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC 2BOY"  
/rissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: PCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
library made from pooled tissue from testis, thymus,  
semilandonous muscle, longissimus muscle, pancreas,  
adrenal, and endometrium."

BASE COUNT 100 a 182 c 158 g 103 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.51e-05 Length: 543  
Score: 86.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 94.12% Mismatches: 0  
Query Match: 95.56% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967-15 (1-17) x BE753801 (1-543)

QY 1 AAlaAysleuLeuProLeuAspLysAspTyrTyValValAArgGIupProGly 17  
|||||  
Db 333 GCCAAGCTGCTGCTCGACCAAGAGTACTGCTGCTCCGCGAGCCAGGC 383

RESULT 7 453 bp mRNA linear EST 21-JAN-1997  
AA023709  
LOCUS  
DEFINITION  
mh77602.f1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA  
clone IMAGE:456986 5' similar to gp:L32955 Mouse protein tyrosine  
kinase (MOUSE);, mRNA sequence.  
AA023709  
VERSION  
AA023709.1 GI:1487624  
KEYWORDS  
EST.  
SOURCE  
house mouse.  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 453)  
Marr,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Thelning,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.

TITLE  
JOURNAL  
COMMENT  
The Washu-HHMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
Washu-HHMI Mouse EST Project  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LIND; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:273874  
Seq primer: -28M13 rev2 from Amersham  
High quality sequence stop: 406.  
Location/Qualifiers  
1. 453  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_image="IMAGE:456986"  
/clone\_lib="Soares mouse placenta 4NDMP13.5 14.5"  
/sex="unknown"  
/rissue\_type="placenta"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Organ: placenta; Vector: pT73D-Pac (Pharmacia)  
with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
1st strand cDNA was primed with a Not I - oligo(dT) primer  
15',  
TGTTACCATCTGTAAGTGGAGCGCCGCGAATTTTATTTTATTTTATTTTATTTT  
T 3'; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M.Patima Bonaldo."

BASE COUNT 79 a 153 c 137 g 84 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.13e-05 Length: 453  
Score: 83.00 Matches: 16  
Percent Similarity: 94.12% Conservative: 0  
Best Local Similarity: 94.12% Mismatches: 1  
Query Match: 92.22% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967-15 (1-17) x AA023709 (1-453)

QY 1 AAlaAysleuLeuProLeuAspLysAspTyrTyValValAArgGIupProGly 17  
|||||  
Db 30 GCTAAGCTGCTGCTCGGGAAGAGTACTGCTGCTCCGCGAGCTGCG 80

RESULT 8 478 bp mRNA linear EST 21-JAN-1997  
AA023670  
LOCUS  
DEFINITION  
mh78606.f1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA  
clone IMAGE:457114 5' similar to gp:L32955 Mouse protein tyrosine  
kinase (MOUSE);, mRNA sequence.  
AA023670  
VERSION  
AA023670.1 GI:1487587  
KEYWORDS  
EST.  
SOURCE  
house mouse.  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 478)  
Marr,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Thelning,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.

TITLE  
JOURNAL  
The Washu-HHMI Mouse EST Project  
Unpublished (1996)

## COMMENT

Contact: Marra M/Mouse EST Project  
 WashU-HHMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:274002

Seq primer: -28m13 rev2 from Amersham  
 High quality sequence stop: 463.

## FEATURES

## source

Location/Qualifiers

1..478

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="IMAGE:457114"

/clone\_lib="Soares mouse placenta 4NbMP13.5 14.5"

/sex="unknown"

/tissue\_type="placenta"

/dev\_stage="adult"

/lab\_host="DH10B"

/note="Organ: placenta; Vector: pT7T3D-Pac (Pharmacia)

with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;

1st strand cDNA was primed with a Not I - oligo(dT) primer

15'

TGTTACCATCTGAGTGGAGCGCCGCAATTTT

T 3'; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M.Fatima Bonaldo."

## BASE COUNT

83 a 160 c 142 g 93 t

## ORIGIN

90 a 171 c 156 g 94 t 1 others

## Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

4,46e-05 Length: 478  
 83.00 Matches: 16  
 94.12% Conserved: 0  
 94.12% Mismatches: 1  
 92.22% Indels: 0  
 Gaps: 0

US-09-397-967-15 (1-17) x AA023670 (1-478)

QY 1 AAlaLysLeuLeuProLeuAspLysAspTyrTyValValArgGluProGly 17

DB 30 GCTAAGCTGCTGCTCCCTGGGAAAGAGACTACTACGTGTCGCGAGCCTGCG 80

RESULT 9

AA510093

LOCUS

DEFINITION

512 bp mRNA linear EST 08-JUL-1997

VS35C07.r1 Soares\_mammary\_gland\_NbMKG Mus musculus cDNA clone

IMAGE:863340.5, similar to gb:L32955 Mouse protein tyrosine kinase

(MOUSE);, mRNA sequence.

AA510093

VERSION

AA510093.1 GI:2247947

KEYWORDS

EST.

SOURCE

house mouse.

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 512)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Thelking,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU-HHMI Mouse EST Project

Unpublished (1996)

CONTACT: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:507428

Seq primer: -28m13 rev2 ET from Amersham  
 High quality sequence stop: 373.

## FEATURES

## source

Location/Qualifiers

1..512

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="IMAGE:863340"

/clone\_lib="Soares\_mammary\_gland\_NbMKG"

/sex="male"

/tissue\_type="mammary gland"

/dev\_stage="4 weeks"

/lab\_host="DH10B"

/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia)

with a modified polylinker; Site\_1: Not I; Site\_2: Eco

RI; 1st strand cDNA was primed with a Not I - oligo(dT)

primer 15'

TGTTACCATCTGAGTGGAGCGCCGCAATGTTT

T 3'; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT7T3 vector.

RNA provided by Dr. Minoru Ko, Wayne State Univ. Library

constructed and normalized by Bento Soares and M.Fatima

Bonaldo."

## BASE COUNT

90 a 171 c 156 g 94 t 1 others

## ORIGIN

90 a 171 c 156 g 94 t 1 others

## Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

4,92e-05 Length: 512  
 83.00 Matches: 16  
 94.12% Conserved: 0  
 94.12% Mismatches: 1  
 92.22% Indels: 0  
 Gaps: 0

US-09-397-967-15 (1-17) x AA510093 (1-512)

QY 1 AAlaLysLeuLeuProLeuAspLysAspTyrTyValValArgGluProGly 17

DB 2 GCTAAGCTGCTGCTCCCTGGGAAAGAGACTACTACGTGTCGCGAGCCTGCG 52

RESULT 10

AA881653

LOCUS

DEFINITION

593 bp mRNA linear EST 26-MAR-1998

VX21C02.r1 Soares\_thymus\_2NbMT Mus musculus cDNA clone

IMAGE:1265090.5, similar to gb:L40172 Mus musculus JAK3 gene,

complete cds (MOUSE);, mRNA sequence.

AA881653

VERSION

AA881653.1 GI:2990963

KEYWORDS

EST.

SOURCE

house mouse.

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 593)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Thelking,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

The WashU-HHMI Mouse EST Project

Unpublished (1996)

CONTACT: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine



COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Plate: LLM10850 row: 9 column: 07  
High quality sequence stop: 709.

FEATURES  
source location/Qualifiers  
1..720  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4925598"  
/lab\_host="NCI CGAP-SG2"  
/note="Organ: salivary gland; Vector: PCMV-SPORT6; Site: 1; NCI; Site\_2: Salt; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 123 a 239 c 229 g 129 t

ORIGIN

Alignment Scores:  
Pred. No.: 7.99e-05 Length: 720  
Score: 83.00 Matches: 16  
Percent Similarity: 94.12% Conservative: 0  
Best Local Similarity: 94.12% Mismatches: 1  
Query Match: 92.22% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967-15 (1-17) x BG87355 (1-720)

QY 1 AlalysleuleuProleuAspTYrTYrValValArgGluProGly 17  
|||||  
Db 229 GCTAAGCTGCTCCCTGGGAAGACTACTAGTGTCGCCGAGCCTGCG 279  
|||||

RESULT 13  
BG870513 907 bp mRNA linear EST 29-MAY-2001  
LOCUS 602791432F1 NCI CGAP-SG2 Mus musculus cDNA clone IMAGE:4922742 5',  
DEFINITION mRNA sequence.  
ACCESSION BG870513  
VERSION BG870513.1 GI:14221053  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 907)  
NIH-MGC <http://mgc.nci.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Plate: LLM10842 row: P column: 07  
High quality sequence stop: 742.

FEATURES  
source location/Qualifiers  
1..907  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4922742"

/clone.lib="NCI CGAP-SG2"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: salivary gland; Vector: PCMV-SPORT6; Site: 1; NCI; Site\_2: Salt; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 166 a 279 c 275 g 185 t

ORIGIN

Alignment Scores:  
Pred. No.: 0.000111 Length: 907  
Score: 83.00 Matches: 16  
Percent Similarity: 94.12% Conservative: 0  
Best Local Similarity: 94.12% Mismatches: 1  
Query Match: 92.22% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967-15 (1-17) x BG870513 (1-907)

QY 1 AlalysleuleuProleuAspTYrTYrValValArgGluProGly 17  
|||||  
Db 366 GCTAAGCTGCTCCCTGGGAAGACTACTAGTGTCGCCGAGCCTGCG 416  
|||||

RESULT 14  
A1981562 592 bp mRNA linear EST 07-MAY-2001  
LOCUS 602791432F1 NCI CGAP-SG2 Mus musculus cDNA clone IMAGE:4922742 5',  
DEFINITION pat.pk0060.a5.f chicken activated T cell cDNA Gallus gallus cDNA  
ACCESSION A1981562  
VERSION A1981562.1 GI:5884590  
KEYWORDS EST.  
SOURCE chicken.  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae; Gallus.  
1 (bases 1 to 592)  
Tirunagaru, V.G., Sofer, L., Cul, J. and Burnside, J.  
An expressed sequence tag database of T-cell-enriched activated chicken splenocytes: sequence analysis of 5251 clones  
Genomics 66 (2), 144-151 (2000)  
JOURNALS 20318616  
MEDLINE  
CONTACT: Joan Burnside  
Molecular Endocrinology  
University of Delaware  
40 Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-1345  
Fax: 302-831-3411  
Email: joan@udel.edu, www.chickest.udel.edu  
Seq primer: 77.

FEATURES  
source location/Qualifiers  
1..592  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
/clone="pat.pk0060.a5.f"  
/clone.lib="chicken activated T cell cDNA"  
/sex="male"  
/cell\_type="Con A-activated splenic T cell"  
/lab\_host="E.coli TOP10 F"  
/note="vector: pCDNA3"

BASE COUNT 113 a 177 c 184 g 109 t

ORIGIN

Alignment Scores:  
Pred. No.: 0.000214 Length: 592  
Score: 80.00 Matches: 15  
Percent Similarity: 94.12% Conservative: 1  
Best Local Similarity: 88.24% Mismatches: 1  
Query Match: 86.89% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967-15 (1-17) x A1981562 (1-592)



GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 28, 2003, 13:46:54 ; Search time 3442.75 Seconds

(without alignments)  
5169.955 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 5860

Sequence: 1 MAPSEPTPLIPORCSLSS.....RPAFATLSPQDPLMRGPG 1099

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-O=/cgn2.1/USPRO.spool/US09397967/cunat_26042003.182313.10728/app.query.fasta.1.1486
-DB=EST_QFMT-fastcap -SUFFIX=first -MINMATCH=0.1 -LOOFCLE=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=biolum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967@cgn2.1.3008.grunat.26042003.182313.10728 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlo:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rnd:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	23.5	971	13	B1411962
2	1358	23.2	1625	11	BC027234
3	1192	20.3	828	13	B1557690
4	1141	19.5	698	14	BM935182
5	1134	19.4	675	12	BF150250
6	1031.5	17.6	907	12	BC670513
7	1028.5	17.6	720	12	BC673355
8	1016	17.3	621	10	AW742526
9	1012	17.3	731	9	AA920299
10	960	16.4	583	10	BB631706
11	935.5	16.0	871	14	BQ709825
12	930	15.9	559	12	BC800853
13	906	15.5	769	9	AJ442008
14	879	15.0	545	12	BE916148
15	874	14.9	539	12	BE913247
16	835.5	14.3	543	12	BE753801
17	831.5	14.2	562	12	BF193178
18	828	14.1	580	10	BB608771
19	824	14.1	557	13	BI338676
20	823	14.0	544	12	BF705690
21	807	13.8	590	14	W48204
22	782.5	13.4	867	9	AA755769
23	781.5	13.3	554	12	BF191625
24	762	13.0	767	13	BI734179
25	761	13.0	763	9	AJ399251
26	759.5	13.0	564	13	BM088287
27	725.5	12.4	540	14	AW505513
28	723	12.3	925	14	BM683553
29	721	12.3	1097	13	BM455276
30	720.5	12.3	526	12	BF191620
31	709.5	12.1	593	9	AA881653
32	705	12.0	501	10	BE623895
33	702	12.0	436	10	AW141115
34	698	11.9	453	9	AA023709
35	695.5	11.9	462	10	AA408832
36	693	11.9	512	9	AA510093
37	693	11.8	478	9	AA023670
38	680.5	11.6	737	9	AU134500
39	666.5	11.4	491	12	BF441858
40	651.5	11.1	923	9	AU133110
41	644.5	11.0	550	10	BE309453
42	638.5	10.9	562	10	AW231037
43	637.5	10.9	472	10	AW785501
44	633.5	10.8	917	12	BG767031
45	628.5	10.7	689	12	BF614255

## ALIGNMENTS

RESULT 1  
B1411962  
LOCUS B1411962 971 bp mRNA EST 14-AUG-2001  
DEFINITION 602966375P1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5121786 5',  
ACCESSION B1411962  
VERSION B1411962.1 GI:15172885  
KEYWORDS EST,  
SOURCE house mouse,  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 971)  
NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL COMMENT

Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LLM11297 row: e column: 19  
 High quality sequence start: 25  
 High quality sequence stop: 785.  
 Location/Qualifiers

## FEATURES

source

```
1. 971
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5121786"
/clone_1ib="NCI-CGAP_Lu33"
/issue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Lung; Vector: pTZ19-3D-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5' TGTTACCATCTGAGTGGAGGCGGCGCTGTTTCTTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTZ193 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
```

BASE COUNT 216 a 274 c 279 g 202 t

ORIGIN

## Alignment Scores:

Pred. No.: 1.43e-129 Length: 971  
 Score: 1380.00 Matches: 287  
 Percent Similarity: 89.30% Conservative: 5  
 Best Local Similarity: 87.77% Mismatches: 27  
 Query Match: 23.55% Indels: 10  
 DB: 13 Gaps: 2

US-09-397-967-16 (1-1099) x B1411962 (1-971)

```

QY 475 ProlysgluysSerAsnleuValValArgArgGlyCysAsnProAlaProAlaPro 494
    |||::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 6 CCAAGGTCAGAGTATTGATGCTGTGGAGAGGGGCTGACCCCGCGCTGCCCCCT 65
QY 495 GlyCysSerProSerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAsp 514
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 66 GCGTGCCTCCCGCTGCTGCTGCGCTGACACAGCTGACCAACAATGCCAAGCGAC 125
QY 515 SerleuGluThrPheHisGluAsnleuGlyHisGlySerPheThrIlePheArgGlySer 534
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 126 AGCTGTGAGTGGCGACGAGAACCTGGGTACACGCTCTTTTACCAGAACTTCGCGGCCGC 185
QY 535 ArgArgGluValValAlaSerGlyGluThrHisAspSerGluValLeuLeuValMetAsp 554
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 186 AGCGGGGAGTGTGATGTGTGACACACATGCTCGGAAGTCCCTCCTGAAGTATGAGAC 245
QY 555 SerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGluVal 574
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 246 TCCAGACATCGAGACTGATGAGAGCTTTCTTGAGAGCCGCAAGCTTGATGAGCAAGTA 305
QY 575 SerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetVal 594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 306 TCCACACCGGACCTGCTGTACTGACAGCGCTGTCATGAGTGGAGACAGATCATGTGTG 365
QY 595 GlnGluPheValTyrLeuGlyAlaAlaIleAspMetTyrLeuArgTylsArgGlyHisLeuVal 614
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 366 CAGGAATTTTGTGTATCTAGAGGACATTCATGACTGCGCAAGCGTGGCCACTGTGTG 425

```

```

QY 615 SerAlaSerTyrPylsLeuGlnValThrTylsGlnLeuAlaTyrAlaLeuAsnTyrLeuGlu 634
DB 426 TCAGCCAGCTGGAAACGACAGGAGGACCAACAGCTGCAATATGCTTAAGTCTTGAG 485
QY 635 AsplysGlyLeuProHisGlyAsnValSerAlaArgTylsValLeuAlaAargGluGly 654
DB 486 GACAAAGGCTTCTCTCAGCGCAAGCTGTACAGACGAGAGGTGCTCGCTGAGAGGG 545
QY 655 GlyAspGlyAsnProPheIleLeuLeuSerAspProGlyValSerProPheValLeu 674
DB 546 GGTGATGGAGATCCACCTTTCATTAACTGATGATCTGCTGTGCACTCCACGTGCTG 605
QY 675 SerleuGluMetLeuThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAla 694
DB 606 AGCTGGAAATGCTCACCAGCAGATACCTGCTGTGCCCCCAATATCTCCAGAGGCT 665
QY 695 GlnThr-LeuGlyLeuGlnAlaAspLysTyrGlyPheGlyAlaThrThrTyrGluValPh 714
DB 666 CAGACACTCTGCTTGGAGGCTGACAGTGGGCTTGGAGCCACCTGGGAAAGTGT 725
QY 714 eGlnArgGlyProAlaHisIleThrSerIleGlu-ProAlaLysGlyLeuLysPheTyrG 734
DB 726 CAGCGGGGACCCGCCAAATCACCCTGCTGAGAGCCGCCAAAAGCTGAAGCTTAG 785
QY 734 LysAspGlnGlyGlnLeuProAlaLeu---LysTyrPheGlnLeuAlaGly-LeuIleTh 752
DB 786 AGGACGAGGAGCAAGTCCCGGCTCTCAATGAGACAGAACTGCGGAGAACTATACAC 845
QY 752 rGlnCys-MetAlaTyrAspProGlyLysArgProSerPheAlaAlaIleLeu-ArgAsp 771
DB 846 ACACTGATGGCTATGAAATCTGCGCGGCGCCCTTATTCGAGATATCTTCAGAGAC 905
QY 772 LeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIlePro 791
DB 906 CTCACGCGCT-ATTACATTGTTTACGAGCTGCTCTA-AAACCCAAACTGGAT---CCC 960
QY 792 SerProArg 794
DB 961 AGTCGAGAGA 969

```

RESULT 2  
 BC027234  
 LOCUS  
 DEFINITION  
 Mus musculus, similar to Janus kinase 3, clone IMAGE:3489805, mRNA.  
 ACCESSION  
 BC027234.1 GI:20071020  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1625)  
 DIRECT SUBMISSION  
 Strausberg, R.  
 Submitted (04-APR-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huijck, S.W., Hale, S.M.,  
 Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
 Richards, S., Gibbs, R.A.



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 35 Row: g Column: 19  
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein  
 This clone has the following problem: Incomplete processing.

## FEATURES

location/Qualifiers  
 1..1625

source  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /map="C57BL/6J"  
 /clone="IMAGE:3489805"  
 /tissue-type="Mammary tumor. MAP-NGF alpha model. 7 months old, gross tissue."  
 /clone\_lib="NCI\_CGAP\_Mam5"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 BASE COUNT 328 a 470 c 483 g 344 t  
 ORIGIN

## Alignment Scores:

Pred. NO.: 6.07e-127 Length: 1625  
 Score: 1358.00 Matches: 281  
 Percent Similarity: 75.59% Conservative: 7  
 Best Local Similarity: 73.75% Mismatches: 11  
 Query Match: 23.17% Indels: 84  
 DB: 11 Gaps: 3

US-09-397-967-16 (1-1099) x BC027234 (1-1625)

Oy 1 Metalprothrombin-like protein 20  
 Db 42 ATGGACCTCCAGTGAAGACACCTGATCCTCAGCGCTTTCAGCTTCATCC 101  
 Oy 21 Serpina1a-like protein 40  
 Db 102 TCAGAGGAGGAGCCGATGCTCCTCCCGGAGACCTGGGCTCCAGGCA 161  
 Oy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuGlyValAlaAla 60  
 Db 162 TTGTATTCCTCTTTGGGAGACTGCTGGGATTTATGTCGACCTCCCAAGGC 221  
 Oy 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSer 80  
 Db 222 TGTGCAATCTGCTGCTTATCATTCGCTTTCGCTGCGCAGAGAGCTTCTTGC 281  
 Oy 81 TrpPheProPheSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
 Db 282 TGGTTCCCAAGGACACATCTTGCATAGAGAGAGCTGACATCAACTTGGCTTAC 341  
 Oy 101 ArgLeuAlaGlyPheTyrPheProAspTrrPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
 Db 342 AGGCTACGCTTTTATTTCCCTACTGCTGGTGGGAGACATGTCACCGCTTTGGGCTG 401  
 Oy 121 ArgValAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLeuPheAlaGln 140  
 Db 402 CCGCAAGATTTTACAGCTCCATCTGCTTACATGTTTGAACAAGCTTTGGTCTG 461  
 Oy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
 Db 462 CACCCAGAGGACCTGCTGAGTGGGCGCCCTCCGCTGGGCTTACGATGAAGAGGAGGGA 521  
 Oy 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnGlnAlaGlnArg 180  
 Db 522 GAGTTCTGAGCTGCTGCTGAGCTTGGCCAGATGCTGAGAGGAGGCCAGGCC 581  
 Oy 181 ProGlyGluLeuLeuLysThrVal 188  
 Db 582 CCAGAGAGAGCTGCTGAAGAGCGT - CAGGTGAGGAGGCTGGGCGACCCCTTCGTTCTTCC 640  
 Oy 188 188

Db 641 TTTGCTGGGATCTAGAGCCCTATGTTTCCCTTACTGGAACATACCTCTTCCA 700  
 Oy 188 188  
 Db 701 ACTTGTGAGCCACCTCTCTGTCAGCTGCCCCGCGAAGGTCCTATGTCTCT 760  
 Oy 189 189  
 Db 761 CGTGTCTCTTGGGCCCCCTCACCACCTAGAGCTGCGGCTGCCCCCCCCCAAGT 820  
 Oy 190 TTTTValaCysLeuProPheSerLeuArgAspValIleGlnGlyGlnAsnThr 209  
 Db 821 TCAAGAAGCTGCTCCGCGCCAGCTGGCGGATGTGATCCAGGCGCAGAACTTCTGACA 880  
 Oy 210 ArgArgArgIleArgArgThrValValLeuAlaLeuLeuProCysGlyValArgLeuProGly 229  
 Db 881 CGCAGGCGCATCCGAGGAGCGGTGTGGGCTGGCG - CGGTGTGTGCTGCTGCGCAAGC 939  
 Oy 230 ArgPro-TyralaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisProAla 249  
 Db 940 CGACCGCTACGCGCTCATGGCCAGATATTTCTGGACCTGGAGCGGCTATCCAGCGGC 999  
 Oy 249 aThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlnGluProGlyLeuLeuArg 269  
 Db 1000 CACCAACGAGACTTCCGTGTGGGCTCCCGGCGCCCGCAGAGAGAGCGGCGCTTCTGCG 1059  
 Oy 269 gValaIaGlyAspAsnGlyIleProTrrPheSerHisAspGluLeuPheGlnThrPhe 289  
 Db 1060 TTGTGGCGGGGAGACCGGATCTCCTGAGCTCCGGGAGCCAGC----- 1102  
 Oy 289 sasPheProGluIleValaAspValSerIle--AsnGlnAlaProArgValGlyProAl 308  
 Db 1103 -----GAGGTGCTGGGCTGGGCTGTGAGAAAGCGGGTCCCGGGAGAGCGCTGGC 1152  
 Oy 308 a 308  
 Db 1153 G 1153

## FEATURES

source

1..826  
 /organism="Mus musculus"  
 /strain="129,C57BL/6J,FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5289949"  
 /clone\_lib="NCI\_CGAP\_Mam3"

		147 a	252 c	240 g	187 t	
US-09-397-967-16 (1-1099) x B1557690 (1-826)						
Alignment Scores:						
Pred. No.:	1,58e-110	Length:	826			
Score:	1192.00	Matches:	249			
Percent Similarity:	95.45%	Conservative:	3			
Best Local Similarity:	94.32%	Mismatches:	9			
Query Match:	20.34%	Indels:	1			
DB:	13	Gaps:	7			
BASE COUNT	147 a	252 c	240 g	187 t		
ORIGIN						
1	MetAlaProPSeSerGluGluThrProLeuLeuLeuProGlnArgSerCysSerLeuSer					
37	ATGGCACTCCAGAGAGAGAGACACCTGTGATCCTCAGGCTTTCAGGCTTCATCC					
21	SerGlnAlaGlyAlaLeuHisValLeuLeuProProAlaGlyProGlyProGlnArg					
97	TCAGAGGAGGAGCCCTGCATGCTCTCTCTCCCGGAGCACTGGGCTCCACAGCA					
41	LeuSerPheSerPheGlyAspTrpLeuAlaGluAspLeuCysValArgAlaAlaLysAla					
157	TTGTATTCTC-TTTGGGACTACTTGGCTGGAGATTATGTTGGTCCAGACTCCAGGCC					
61	CysGlyLeuLeuProValTyrHisSerLeuPheAlaLeuAlaThGlnAspPheSerCys					
216	TGTGGCATTCCGCTGTTTATTCATTCGCTTTCCGCTCGGCCATGAGAGACTTCTTGC					
81	TrpPheProProSerHisIlePheCysIleGluAspValAspTrpGlnValLeuValTyr					
276	TGGTTTCCCCCAAGGACACATCTTGCATAGAGGAGCTGGACACATCTTGGCTTAC					
101	ArgLeuAcGpHeTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu					
336	AGGCTACGCTTATTATTCCTCAGCTGGTTTGGGCTGGAGACATGTCACCGCTT-GGGCTG					
121	ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln					
395	CGCAAAAGATTG-ACCAGAGGCCATCTTGACTTACATTACTGTTTAGAACATCTCTTCTCAG					
141	HisArgSerAspLeuValSerGlyArgLeuProValIleGlyLeuSerMetLysGluGlnGly					
454	CACGCCACTGACCTGTGTAGTGGGCGCTCCCGGAGGCCCTTAGCATGAAGAGAGAGGA					
161	GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnAlaGlnArg					
514	GAGTTCCTGAGAGCTGGCGGTGTGTGACTTGGCCAGATGGCTCGTAGACAGGCCACAGGC					
181	ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp					
574	CCAGAGAGAGCTGTGTAACAGCGTACGTTAAAGCGCTGTGCCGCCACAGCCTGGCGCAT					
201	ValIleGlnGlyGlnAspPheValThrArgArgArg-1LeuArgArgThrValValLeuAl					
634	GTAATCCAGGGCGAACTCTGTGTGACAGCAGGCAAGCATCCGCAAGACCGTGTCTTGGC					
220	AlaLeuLeuProCysGlyArgLeuProGlyArgProGlyAlaLeuMetAlaLysTyrIleLe					
694	GCTGG-CCGTTGTGTGCTTTCAGGCGCCAGCCAGTACCGCTCATGGCCAGTATATCT					
240	AspLeuGluAlaArgLeuHisPro-AlaAlaThrThrGluThrPheArgValGlyLeuProG					
753	GGACTGGAAACGATACCTCCAGAGCGGCACACAGAGA---CTCCGTTTGGGGCTCCCGG					

DB	810	GCGCCAGG	817
RESULT 4			
LOCUS	BM935182		
DEFINITION	UI-M-BH3-agz-f-08-0-UI r1 NIH_BMAP_M_S4 Mus musculus cDNA clone	698 bp	mRNA linear EST 13-MAR-2002
ACCESSION	UI-M-BH3-agz-f-08-0-UI 5', mRNA sequence.		
VERSION	BM935182.1	GI:19394334	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 698)		
	Bonaldo,M.F., Lennon,G. and Soares,M.B.		
	Normalization and subtraction: two approaches to facilitate gene		
	discovery		
JOURNAL	Genome Res. 6 (9), 791-806 (1996)		
MEDLINE	97044477		
COMMENT	Contact: Chin, H National Institute of Mental Health 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA Tel: 301 443 706 Fax: 301 443 9890 Email: MEST@mail.nih.gov cDNA library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).		
FEATURES	Seq primer: M13 REVERSE.		
Source	Location/Qualifiers		
	1..698		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="UI-M-BH3-agz-f-08-0-UI"		
	/clone_1lb="NIH_BMAP_M_S4"		
	/dev_stage="27-32 days"		
	/lab_host="DH10B (Life Technologies)"		
	/note="Vector: pUT3D-Pac (Pharmacia) with a modified		
	polylinker. Site.1: Not I; Site.2: Eco RI; The		
	NIH_BMAP_M_S4 library is a subtracted library of a series,		
	ultimately derived from a mixture of individually tagged		
	normalized libraries from ten regions of the mouse brain		
	(cerebellum, brain stems, olfactory bulbs, hypothalamus,		
	cortex, amygdala, basal ganglia, pineal gland, striatum,		
	hippocampus) after a series of subtractions to reduce the		
	representation of cDNAs from which ESTs had already been		
	generated. The following serially substracted libraries		
	were generated in this process: NIH_BMAP_M_S1,		
	NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,		
	NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library		
	(NIH_BMAP_M_S4) was constructed as follows: Peramplified		
	cDNA inserts from NIH_BMAP_M_S3.3', NIH_BMAP_M_S3.2', and		
	NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived		
	was used as a driver in a hybridization with a pool of		
	the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1		
	libraries in the form of single-stranded circles. The		
	remaining single-stranded circles (subtracted library)		
	was purified by hydroxyapatite column chromatography,		
	converted to double-stranded circles and electroporated		
	into DH10B bacteria (Life technologies) to generate the		
	NIH_BMAP_M_S4 library. This procedure has been previously		
	described (Bonaldo, Lennon and Soares, Genome Research		
	6:791-806, 1996)".		

Alignment Scores: 1.88e-105 Length: 698  
 Pred. No.: 1141.00 Matches: 225  
 Score: 96.15% Conservative: 0  
 Percent Similarity: 96.15% Mismatches: 6  
 Best Local Similarity: 96.15% Gaps: 2  
 Query Match: 19.47% Indels: 4  
 DB: 14

US-09-397-967-16 (1-1099) x BM935182 (1-698)

OY 792 SerProTgAspGluLeuGlyValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAla 811  
 |||||  
 DB 3 AGTCCCTCGAGATGAGCTGTGC---GGTGGCGCCCAAGCTCTATGCTGCCAGAGACCCCGCC 59  
 |||||

OY 812 IllePheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlnPheGlySer 831  
 |||||  
 DB 60 AATATTCGAGAGAGACACCTTAAGTACATCTCTTGGCGCAAGGCAACTTTGGCAGC 119  
 |||||

OY 832 ValGluLeuGlyArgTyrAspProLeuGlnYAspAsnThrGlyProLeuValAlaValLys 851  
 |||||  
 DB 120 GTGGAGCTGTGCCCTATGACCCCTGGGGGACATACGGGACCCCTGGTGGCAGTGAA 179  
 |||||

OY 852 GlnLeuGlnHisSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeu 871  
 |||||  
 DB 180 CAGCTACAGACACAGCGGGCCAGACAGAGAGACTTCACAGCGGAGATCAGATCCTT 239  
 |||||

OY 872 LysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArg 891  
 |||||  
 DB 240 AAGGCTGTGCACAGGACATTCATCGTCAAGTACCGGGAGTACAGCTATGGCCAGGCTCC 299  
 |||||

OY 892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGln 911  
 |||||  
 DB 300 CAGACCTCGGCTGTGTGATGAGTACCTGCCAGCGGCTGGCTGGAGATCTCCGAGC 359  
 |||||

OY 912 ArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLys 930  
 |||||  
 DB 360 CCCCATCGCGCCGCTGCACACCGCCCTACTGCTGTGCTGGCGCAGATCTGCAG 419  
 |||||

OY 931 GlyMetGluTyrLeuGlnValArgArgCysValHisArgAspLeuAlaIleArgAsnIle 950  
 |||||  
 DB 420 GGCATGCACTACCTGGGTGGCGCCGCTGCTGACACCGTACCTGGCTGGCGCAACATC 479  
 |||||

OY 951 LeuValGluSerGlnAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro 970  
 |||||  
 DB 480 TTGGTGGAGAGCGAGGCTCATGTGAAGATCGCGGACTTGCCTCGTAAAGCTGCGCC 539  
 |||||

OY 971 LeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIle-PheTrpTyrAl 990  
 |||||  
 DB 540 CTGGGAAAGGACTACTACTGTGCTGGCGGAGCTGGCCAAAGCCCACTTNTTGGTATGC 599  
 |||||

OY 990 AProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyVal 1010  
 |||||  
 DB 600 CCCGAGATCCCTATCTGACAC-ATCTTCTCNCGCCCAATGTGACGTGGAGCTTCGAGT 658  
 |||||

OY 1010 lValLeuTyrGluLeuPheThrTyrCysAspLysSerCys 1023  
 |||||  
 DB 659 GGTGTTGTACGAGCTCTTCACTACTGCGACAAGAGCTGC 698  
 |||||

RESULT 5  
 BF150250 675 bp mRNA linear EST 29-DEC-2000  
 LOCUS BF150250  
 DEFINITION uy83e09.y1 NCI\_CGAP\_Mam5 Mus musculus cDNA IMAGE:3666184 5'  
 similar to TR:p97423 p97423 JANUS KINASE 3 ;, mRNA sequence.  
 accession BF150250  
 version BF150250.1 GI:11031645  
 keywords EST.  
 source house mouse.  
 organism Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 675)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Other-ESTs: uy83e09.x1  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Lotnar Hemmighausen Ph.D., Robin Humphreys  
 Tissue Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone Distribution: NCI-CGAP clone distribution Information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 image.llnl.gov/image/html/resources.shtml  
 MGI:1426952

Seq primer: -40RP from Gibco  
 High quality sequence stop: 390.  
 Location/Qualifiers  
 1..675  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone\_image="366184"  
 /clone\_lib="NCI\_CGAP\_Mam5"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="7 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-Sport6; Site:1; Salt:  
 Site:2; Note: Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lotnar Hemmighausen/Robin Humphreys,  
 NIH"

BASE COUNT 138 a 215 c 197 g 123 t 2 others  
 ORIGIN

Alignment Scores: 9.15e-105 Length: 675  
 Pred. No.: 1134.00 Matches: 217  
 Score: 96.02% Conservative: 0  
 Percent Similarity: 96.02% Mismatches: 7  
 Best Local Similarity: 96.02% Indels: 2  
 Query Match: 19.35% Gaps: 2  
 DB: 12

US-09-397-967-16 (1-1099) x BF150250 (1-675)

OY 705 GlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeu 724  
 |||||  
 DB 1 GGCTTTGGAGCCACACGCTGGAGGTCTTCACGGGGGACCCGCCACATCAGCTGCTG 60  
 |||||

OY 725 GluProAlaLysLysLeuLysPheTyrGluAspGlnGlyLeuProAlaLeuLysTyr 744  
 |||||  
 DB 61 GAGCCCGCCAAAAGCTGAAGTCTATGAGAGACAGGACAGCTGCCGCTCAAAATGG 120  
 |||||

OY 745 ThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSer 764  
 |||||  
 DB 121 ACAGAACTGCGCGGACTTATCACACAGTCAATGCGCTATGATCCCGCGGCCCTCC 180  
 |||||

OY 765 PheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSer 784  
 |||||  
 DB 181 TTCCGAGCTATCTTCAGAACTCTCAACGCGCTTCATTACATCAGATTAACGAGCTCTCTCA 240  
 |||||

OY 785 AspProThrProGlyIleProSerProArgAspGluLeuGlyValAlaGlyAlaGlnLeu 804  
 |||||  
 DB 241 GACCCACACACTGGGATCCGAGTCTCGATGATGACTGTGC---GGTGGCGCCAGCTC 297  
 |||||

OY 805 TyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyrIleSerLeuLeu 824  
 |||||  
 DB 298 TATGCTGTCCAGACCCGCCCATATTCGAGGAGACACCTTAAGTACATCTCTTGGCTG 357  
 |||||

OY 825 GlyLysGlyAsnPheGlySerValGluLeuGlyArgTyrAspProLeuGlnYAspAsnThr 844  
 |||||  
 DB 358 GGCMAAGGCACTTTGGCAGCGGTGGAGCTGTGCCGTATGACCCCTGGGGGACATATACG 417  
 |||||

QY 845 GLYPLeuValAlaValLysGlnLeuGlnHisSerValProAspGlnGlnArgAspPhe 864  
 |||||||  
 Db 418 GCACCCCTGGCGAGTGAACACCTCAGACAGCGGCGCACACAGCAGGAGACTTC 477  
 QY 865 GlnArgGlnGlnGlnLeuLysAlaLeuHisSerAspPheLeuValLysTyrArgGly 884  
 |||||||  
 Db 478 CAGGGAGATCGATCGCTTAAGCTCTGACAGCGATTCATCGTCAATACCGGGGA 537  
 QY 885 ValSerTyrGlnProGlnArgGlnSerLeuArgLeuValMetGlnTyrLeuProSerGly 904  
 |||||||  
 Db 538 GTGAGCATATGGGCGCAGCTCGCAGAGCTGTGTGATGATGAGTACCTCCAGCGGC 597  
 QY 905 CysLeuArgAspLeuLeuGlnArgHisArgGly---LeuHisTyrAspArgLeuLeu 923  
 |||||||  
 Db 598 TGGCTGGAGACTTTCGACAGCGCCATCGCGCGCTGACACCGAGCGGCTACTGCTG 657  
 QY 924 PheAlaTyrGlnIleCys 929  
 |||||||  
 Db 658 TTGCCTGGCAGATCTGC 675  
 RESULT 6  
 BG870513 907 bp mRNA linear EST 29-MAY-2001  
 LOCUS 602791432F1\_NCL\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4922742 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG870513 GI:14221053  
 VERSION BG870513.1  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 907)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L1AM10842 row: P column: 07  
 High quality sequence stop: 742.  
 FEATURES  
 source Location/Qualifiers  
 1. 907  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4922742"  
 /clone\_lib="NCL CGAP\_SG2"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;  
 Ntl; Site: 2; Salt: Cloned unidirectionally. Primer: Oligo  
 dt. Average insert size 1.3 kb. Constructed by Life  
 Technologies. Note: this is a NCL CGAP Library."  
 BASE COUNT 166 a 279 c 275 g 185 t 2 others  
 ORIGIN  
 Alignment Scores: Length: 907  
 Pred. No.: 4.27e-94  
 Score: 1031.50 Matches: 216  
 Percent Similarity: 88.93% Conservative: 1  
 Best Local Similarity: 88.52% Mismatches: 23  
 Query Match: 17.60% Indels: 7  
 DB: 12 Gaps: 3  
 US-09-397-967-16 (1-1099) x BG870513 (1-907)  
 QY 858 ProAspGlnGlnArgAspPheGlnArgGlnIleLeuLysAlaLeuHisSerAsp 877

|||||||  
 Db 39 CCAGACAGGAGAGGAGCTTCACAGCGGAGATTCAGATCTTAAGGCTTCACAGCGAC 98  
 QY 878 PheIleValLysTyrArgGlyValSerTyrGlnProGlnArgGlnSerLeuArgLeuVal 897  
 |||||||  
 Db 99 TTCATGTCAGATACCGGGGAGTCACTATGGCCAGGTCCCGCAGAGCTTCGGTGGTG 158  
 QY 898 MetGlnTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly---Leu 916  
 |||||||  
 Db 159 ATGAGATACCTGCCAGGGGCTGCTCGAGAGATTCCTCCAGCGGCCATCCGGCGGCTG 218  
 QY 917 HisThrAspArgLeuLeuPheAlaTyrGlnIleCysGlyGlyMetGlnTyrLeuGly 936  
 |||||||  
 Db 219 CACACCGACCGCTTACGCTGCTTGGTGGAGATCTGCAGAGGCGCATGAGTACTCGGCT 278  
 QY 937 AlaArgArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValGlnSerGlnAla 956  
 |||||||  
 Db 279 GCGGCGCGCTGCTGACACCGAGACCTGGCTGGCAGACATCTGTGGAGAGGAGGCT 338  
 QY 957 HisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyr 976  
 |||||||  
 Db 339 CATGTGAAGATCGCGGAGCTTCGCTCGTACGCTGCTGCCCTGGGAAAGACTACTAC 398  
 QY 977 ValValArgGlnProGlnGlnSerProIlePheTyrAlaProGlnSerLeuSerAsp 996  
 |||||||  
 Db 399 GTGGTCGCGCGAGCTGCGCAAGGCCCATCTTTGGTATGCCCGGAGAGTCTATCTGAC 458  
 QY 997 AsnIlePheSerArgGlnSerAspValTyrSerPheGlyValIleLeuTyrGlnLeuPhe 1016  
 |||||||  
 Db 459 AACATCTTCTCCGCCATCTGAGACCTGTGAGCTTCGGAGTGGTGTGACAGCTTTC 518  
 QY 1017 ThrTyrCysAspLysSerCysSerProSerAlaGlnPheLeuArgMetMetGlyProGln 1036  
 |||||||  
 Db 519 AC-TACTGCGCAGAGACTGACGCCATCCGCTAGTTCCTGACACATGATGGGCGCTGAG 577  
 QY 1037 ArgGlnGlyProProLeuCysArgLeuGlnIleLeuAlaGlnGlyArgArgLeuPro 1056  
 |||||||  
 Db 578 CGTGAAGAGACCCCGCTTCTGGCTC-CTGGAGCTGCTGGCAGAGGCGCGAC-CTCCCA 635  
 QY 1057 ProProThrThrCysProThrGlnValGlnGlnLeuMetGlnLeuCysTyrPalaProGln 1076  
 |||||||  
 Db 636 CATCTTCCACTGCC---CACGAGGTTCAAGAGCTCATGACGCTGTCTGGGCCCCAGC 692  
 QY 1077 ProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTyrArgGly 1096  
 |||||||  
 Db 693 CCG---ACGAGCGGCGAGCTTCGAGACCTGAGCCCGCAGCTGACGCGGCTGTGAGTGA 749  
 QY 1097 -ArgProGly 1099  
 |||||||  
 Db 750 AAGACCGGG 759  
 RESULT 7  
 BG873355 720 bp mRNA linear EST 29-MAY-2001  
 LOCUS 602794328F1\_NCL\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4925598 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG873355  
 VERSION BG873355.1 GI:14223895  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 720)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:

http://image.lnl.gov

Plate: LAM0850 row: g column: 07

High quality sequence stop: 709.

## FEATURES

Source

1. 720

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4925598"

/lab\_host="NCI-CGAP\_SG2"

/note="Organ: salivary gland; Vector: PCMV-SPORE6; Site: 1;

NCI; Site\_2: Salt; Cloned unidirectionally. Primer: Oligo

dr. Average insert size 1.3 kb. Constructed by Life

Technologies. Note: this is a NCI-CGAP Library."

BASE COUNT 123 a 229 c 229 g 129 t

ORIGIN

## Alignment Scores:

Pred. No.: 5,81e-94 Length: 720  
Score: 1028.50 Matches: 201  
Percent Similarity: 95.71% Conservative: 0  
Best Local Similarity: 95.71% Mismatches: 8  
Query Match: 17.55% Indels: 3  
DB: 12 Gaps: 1

US-09-397-967-16 (1-1099) x BG873355 (1-720)

QY 891 ArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeu 910

DB 1 CGCGAGGCGTGGGTTGGATGGATGAGTACTCCGAGCGGCTGCGAGACTTCCTG 60

QY 911 GlnArgHisArgGly---LeuHisThrAspArgLeuLeuLeuLeuLeuLeuLeu 929

DB 61 CACGGCCATGGCGCGCGCCGACACCGACCGCCCTACTGCTGCTGGCAGATCTGC 120

QY 930 LysGlyMetCysLutyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaArgAsn 949

DB 121 AAGGGATGAGTACCTGGCTGGCGCGCGCTGCTGACCTGCTGGCGCGCAC 180

QY 950 IleLeuValGlnSerGlnAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeu 969

DB 181 ATCTTGTGAGAGCGAGCTCTGTGAAGATCGCGACTTCGCTGAGTGGCTG 240

QY 970 ProLeuGlyLysAspTyrTyrValAlaArgGluProGlyGlnSerProIlePheTyr 989

DB 241 CCCCTGGGAAGAGTACTAGTGGTGGCGGAGCTGGCCAAAGCCCATCTTTGCTAT 300

QY 990 AlaProGlnSerLeuSerAspAsnIlePheSerArgGlnSerAspValIrrSerPheGly 1009

DB 301 GCGCCGAGTCCCTATCTACACATCTTCCCGCAATCTGACCTGTGAGACTTCGGA 360

QY 1010 ValValLeuTyrGlnLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPhe 1029

DB 361 GTGGTGTGTACAGCTTTCACCTACTCGACAGAGCTGACGCCATCGCTGAGTTC 420

QY 1030 LeuArgMetMetCysProGluArgGluGlyProProLeuGlyCysArgLeuGluLeu 1049

DB 421 CTGGCGATGATGGGCTGAGCGTGAAGAGCCCGCTGCGCGCTCTGGAGCTGCTG 480

QY 1050 AlaGlnGlyArgArgLeuProProProProThrCysProThrGlyValGlnGluMet 1069

DB 481 GCGAGGCGCGAGCGCTCCACCACTCCACCTGCGCCACGAGGTTGAGAGCTCATG 540

QY 1070 GlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeuSerProGln 1089

DB 541 CAGCTGTGCTGGGCGCGCGCGCA-GACCGGCGACGCTTGG-ACCTGAGCCCGCAG 598

QY 1090 LeuAspProLeuTyrArgGlyArgProGly 1099

DB 599 CTGGAGCGCGCTGTGGCTGGAAGACCCGGA 628

RESULT 8  
AW742526  
LOCUS  
DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: Robert Strausberg, Ph.D.

Email: cgabs-remail.nhl.gov

This clone is available royalty-free through LINL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

Seq primer: -40RP from Gibco

High quality sequence stop: 471.

Location/Qualifiers

1. 621

/organism="Mus musculus"

/strain="C3H x 101 F1"

/db\_xref="taxon:10090"

/clone="IMAGE:2780302"

/lab\_host="Soares\_mouse\_NMIE"

/sex="male"

/dev\_stage="newborn"

/lab\_host="DH10B"

/note="Organ: inner ear, 170 pooled; Vector: pRT3D-Pac;

Site\_1: Not 1; Site\_2: Eco RI; 1st strand cDNA was primed

with a Not I - oligo(dT) primer (5'

TGTTCACATCTGAGTGGAGGCGCGCCGACCTTTTCTTTTCTTTTCTTTT

3'); double-stranded cDNA was ligated to Eco RI adaptors

(pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified pRT3D vector. Library

is normalized, and was constructed and donated by Bento

Soares and M. Fatima Bonaldo (University of Iowa) and R.

Hardisty, A. Varela-Carver, P. Mburu and S.D.M. Brown (MRC

UK Mouse Genome Centre and Mammalian Genetics Unit,

Harwell, UK)."

BASE COUNT 124 a 197 c 182 g 118 t

ORIGIN

Alignment Scores:

Pred. No.: 8.47e-93 Length: 621  
Score: 1016.00 Matches: 196  
Percent Similarity: 95.17% Conservative: 1  
Best Local Similarity: 94.69% Mismatches: 8  
Query Match: 17.34% Indels: 2  
DB: 10 Gaps: 2

US-09-397-967-16 (1-1099) x AW742526 (1-621)

QY 741 AlaLeuLysTrpThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760

DB 3 GCTCTCTATATGACAGACTGGCGGACTTATGACACAGCTCATGATGATCTGGC 62

QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780

DB 63 CGGCGCCCTCTCTCGAGCTATCTCAGAGACCTCAACGCGCTCATTAATCAAGATTAC 122

QY 781 GlnLeuLeuSerAspProThrProGlyTyrLeuProSerProAlaGlyLeuCysValAla 800

DB 123 GAGCTCTCTCAGACCCACACACCTGGCATCCGAGTCTTCAGATGAGACTGTC---GCT 179

QY 801 GlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluAlaArgHisLeuLysTyr 820

```

|||||
Db 180 GGGCCAGCTATGCTCCAGACCCGACCATATTCAGAGAGACACCTTAAGTAC 239
Oy 821 11SerleuenglylysglyAsnPhelglySerValgluLeucysArgTyraSPProleu 840
Db 240 ATCTCTTGTGGCAGGCGCACTTGTGGCAGCGTGTAGTGTCCCTATGACCCCTTG 259
Oy 841 G1yAspAsnThrGlyProleuValAlaValIysGlnleuGlnHisSerValProAspGln 860
Db 300 GGGGACATACGGGAGCCCTGTGGCAGTGAACACCTACAGCAGCGGCGGACAGCAG 359
Oy 861 G1nArgAspPhelGlnArgGluIleGlnIleleuYsa1aleuHisSerAspPheIleVal 880
Db 360 CAGAGGAGCTTCCAGCGGAGATTACATCTTAAGCTCTGCACAGCACTTATGTC 419
Oy 881 LysTyraArgGlyValSerTyrglyProGlyArgGlnSerleuArgleuValMetGluTyr 900
Db 420 AAGTACCGGGAGTACAGTATGGCCAAAGTCCAGAGCTCGCGTTGTGTGATGAGTAC 479
Oy 901 LeuProSerGlyCysleuArgAspIleuGlnArgHisArgGly---leuHisThrAsp 919
Db 480 CTGCCACGGCTCTCGTGGAGACTCTCGACGCGCATCGCGCGCTTGCACACCGAC 539
Oy 920 ArgleuLeuPhelAlaTrpGlnIleCysIysGlyMetGluTyrLeuGlyAlaArgArg 939
Db 540 CGGCTACTGCTGTCGCTGGCAGATCTGCAGAGCATGAGTACTGTGTGCGCGCGC 599
Oy 940 CysValHisArgAspLeuAla 946
Db 600 TGCCTACACCTGACCTGGCT 620

RESULT 9
AA920299 731 bp mRNA linear EST 20-APR-1998
LOCUS AA920299
DEFINITION v93902.1 Soares_thymus_2nbmt Mus musculus cDNA clone
IMAGE:1282802 5' similar to gb:140172 Mus musculus JAK3 gene,
complete cds (MOUSE);, mRNA sequence.
VERSION AA920299
KEYWORDS AA920299.1 GI:3067078
SOURCE EST.
ORGANISM house mouse.
MUS musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
1 (bases 1 to 731)
Marras, S., Kucaba, T., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 215.
Location/Qualifiers
1..731
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1282802"
/clone_1id="Soares_thymus_2nbmt"
/sex="male"
/tissue_type="Thymus"

```

```

/dev_stage="4 weeks"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 167 a 209 c 214 g 141 t
ORIGIN
Alignment Scores:
Pred. No.: 2,87e-92 Length: 731
Score: 1012.00 Matches: 213
Percent Similarity: 90.42% Conservative: 4
Best Local Similarity: 88.75% Mismatches: 21
Query Match: 17.27% Indels: 1
DB: Gaps: 6

US-09-397-967-16 (1-1099) x AA920299 (1-731)
Oy 660 ProPheIleLysleuSerAspProGlyValSerProThrValleuSerleuGluMetleu 679
Db 9 CCTTTCATTAAGCTGAGTGTGATCTCGTGTGATCTCCACTGTGTGTGAGAAATGCTC 68
Oy 680 ThrAspArgIleProTrpValAlaProGluCysleuGlnIleGlnIleThrCysLeu 699
Db 69 ACCACAGAAATACCTCGTGGTGGCCCGCAATGTCCTCAGAGAGCTCAGACACTCGCTTG 128
Oy 700 G1uAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAla 719
Db 129 GAGCTCAGCAAGTGGGCTTGTGGAGCCACACGCGGAGGTGTTCACCGGGGACCCG-- 186
Oy 720 HisIleThrSerleuGluProAlaLysIleuLysIleuLysIleuLysIleuLysIleu 739
Db 187 CACATCACCTTCGTCGAGCGCCGCAAAAGCTGAAGTTCTATGAGACACAGGACACTG 246
Oy 740 ProAlaLeuLysTrpTrpGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759
Db 247 CCCGCTCTCAATATGACAGAACTGGCGGACTTTCACAGAGTCAGCGCTATGATTTCT 306
Oy 760 G1yArgArgProSerPheArgAlaIleleuArgAspLeuAsnGlyLeuIleThrSerAsp 779
Db 307 GCGCGGCGCCCTCTTCGAGTAACTCTCAGAGACCTCAACGGCTCATTTACATCAGAT 366
Oy 780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVal 799
Db 367 TACGAGCTCTCTCAGACCCCAAAACCTGGCATCCCGAGTCTCGAGATGAGCTGTGC--- 423
Oy 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLys 819
Db 424 GGTGGGCGCCAGCTCTATAGCTCCGACGAGACCCCGCATATTCAGAGAGACACCTTAAG 483
Oy 820 TyrIleSerleuLeuGlyLysGlyAsnPhelglySerValgluLeuCysArgTyraSPPro 839
Db 484 TACATCTCTTTTGTCT-GGGAGCGGAACTTGGCAGCGTGGAGCTGTGGCTATGACCCC 542
Oy 840 LeuGlyAspAsnThrGlyProleuValAlaValIysGlnleuGlnHisSerValProAs 859
Db 543 CTGGGAGCAATACCGGAGCCCTGTGTGCAAGTGAAGACAGTACAGCAGCGGAGCAGA 602
Oy 859 pGlnGlnArgAspPhelGlnArgGluIleGlnIleleuYsa1aleuHisSerAspPheI 879
Db 603 CCAGCAAAAGAGCTTCCAGCGAGATTCAGATTCCTAAGCTCTGCACACGAGATTCAT 662
Oy 879 eValIysTyraArgGlyValSerTyrglyProGlyArgGlnSerleuArgleuValMet 898
Db 663 GCTCAAGTACCGGGAATCAG-TATGGGCGCAGGTGCCAAGC-CTCGGTTTGATATG 718

```

RESULT 10  
LOCUS BB631706  
DEFINITION BB631706 RIKEN full-length enriched, 16 days neonate thymus Mus musculus cDNA clone A130091E14 5', mRNA sequence.  
ACCESSION BB631706  
VERSION BB631706.1 GI:16468415  
KEYWORDS EST.  
SOURCE mouse mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 583)  
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shitaki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
UNPUBLISHED (2001)  
CONTACT: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa. 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.go.jp/  
URL: http://genome.resc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
waghi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsuhi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamana,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (http://genome.resc.riken.go.jp) for further details.  
e mouse tissues.

FEATURES  
Source  
Location/Qualifiers  
1..583  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="A130091E14"  
/clone\_lib="RIKEN full-length enriched, 16 days neonate thymus"  
/tissue\_type="thymus"  
/dev\_stage="16 days neonate"  
/lab\_host="DH10B"  
/note="Site\_1: SalI; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'

BASE COUNT 106 a 171 c 156 g 150 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,88e-87  
Score: 960.00  
Percent Similarity: 98.94e  
Best Local Similarity: 97.34e  
Query Match: 16,38e  
DB: 10  
Gaps: 0.  
US-09-397-967-16 (1-1099) x BB631706 (1-583)  
QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSer 20  
DB 22 ATGGACACCTCCAGAGTGGAGACACCTCGATCCCTGAGCCTCTGATCC 81  
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40  
DB 82 TAGAGGACGAGGAGCCCTGATGTCCTCCCTCCCGGAGCCTGGGCTCCAGGCA 141  
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuValAlaAlaAlaAla 60  
DB 142 TTGTATCTCTCTTTGGGAGCTACTGGCTGGAGATTATGTCGAGCTCCAGGCC 201  
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
DB 202 TGTGCATCTCTCCGTTATCATTCGCTTTTCCTCGCCAGCAGACGACTCTCTTGC 261  
QY 81 TrpPheProPheSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
DB 262 TGGTTCCTCCCAAGCCACATCTTCTGATAGGAGCGGACACCAAGCTTGGTCTAC 321  
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluTyrCysHisArgPheGlyLeu 120  
DB 322 AGGCTACGCTTTTATTTCCCTGACGCTTGGCTGGAGACATGTCACCGCTTGGGCTG 381  
QY 121 ArgLysAspPheThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
DB 382 CGCAAGATTGGACACGAGCCATCTTGACATCATTAGAAACATCTCTTGGCTCAG 441  
QY 141 HisArgSerAspLeuValSerGlyAlaGluProValGlyLeuSerMetLysGluGlnGly 160  
DB 442 CACCCGATGACCTGGTAGAGGGGCGCTCCGGTGGGCTTTAACTAGAGACACGAGCA 501  
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
DB 502 AAGTTCCGAGACCTGGCGCTGCTGAGCTGGCCCAAGTGCCTGTAGACAGCCACCGC 561  
QY 181 Pro-GlyGluLeuLeuysThr 187  
DB 562 CCAAGGAGAGCTGCTGAAGACA 583  
RESULT 11  
LOCUS B0709825  
DEFINITION B0709825 871 bp mRNA linear EST 16-JUL-2002  
AGENCOURT-84774895 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:6301440  
ACCESSION B0709825  
VERSION B0709825.1 GI:21848724  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens





Matches:	179
Conservative:	0
Mismatches:	7
Indels:	1
Gaps:	0

```

/db_xref="taxon:9031"
/clone="17j3r1"
/clone_lib="dkfz426"
/tissue_type="Bursa of Fabricius"
/dev_type="bursal lymphocyte"
/cell_stage="2-3 weeks old"
222 c 262 g 134 t 1 others

```

Alignment Scores:			
Pred. No.:	2e-81	Length:	769
Score:	906.00	Matches:	182
Percent Similarity:	81.27%	Conservative:	22
Best Local Similarity:	72.51%	Mismatches:	42
Query Match:	15.46%	Indels:	5
DB:	9	Gaps:	3
US-09-397-967-16 (1-1099) x AJ424208 (1-769)			
QY	770	ArgAspLeuAnGlyLeuIleIleTrnSerAspTyrGluLeuSerAspProthProGly	789
DB	1	CGGCACATCAACACCTCATCTCTCCGACTACGACTGCTCTCAAGACTGCACCGGC	60
QY	790	IleProSerProArgAspGlnLeuGlyCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAsp	809
DB	61	GATGTACGCGTGGGG---GAGAGCTCTGTGGGGTACGAGACAGCTGGCGGGCGGCACGGC	117
QY	810	ProAlaIlePheGluGluArgIleHisLeuLysTyrIleSerLeuGlnGlyLysGlnPhe	829
DB	118	CCGGCTCAGTTCGAGGAGGACCACTCAAGTACATCTCACTGCTGGGCAAGGCAACTTT	177
QY	830	GlySerValGlnLeuGlyCysArgTyrAspProLeuGlnLysAspThrGlyProLeuValAla	849
DB	178	GGGAGCTGAGCTGTGCGCTACGACCGCGTGGGTGACAGCACGGGTAGCTGGGGCC	237
QY	850	ValLysGlnLeuGlnHisSerValProAspGlnGlnArgAspPheGlnArgGlnIleGln	869
DB	238	GTCAGAACGCTGCAGCAGAGATTCGGGCAAGAGACCTGCAGAGACTTTCAGAGGAGATCCAG	297
QY	870	IleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyLysTyrGlyPro	889
DB	298	ATCTCGACATCGCTGCAGCAGCACTTCACTGTCACATACCGGGGCTCTGCTACAGCCGT	357
QY	890	GlyArgGlnSerLeuArgLeuValIleMetGlyTyrLeuProSerGlyCysLeuArgAspLeu	909
DB	358	GGGCGCGGGGGCTCGCGTTGGTGGATGAGATATCTGCTGCAGCGCTGCTGAGGGACTAC	417
QY	910	LeuGlnArg-----HisArgGlyLeuHisThrAspArgLeuLeuLeuPheIleArgGln	927
DB	418	CTGCAAAAGAACACAGCACGCC---CTGAGACACCGGACCGCTGGTGTCTACAGCGTGGAG	474
QY	928	IleCysLysGlyMetGlyTyrLeuGlnGlyAlaArgArgCysValHisArgAspLeuAla	947
DB	475	ATCTGCAGAGGGCATGGATGATCACTGGGGGGGACCGCTGGCTCACCGCACTTGGCCACG	534
QY	948	ArgAsnIleLeuValGlnSerGlnAlaHisValLysIleAlaAspPheGlyLeuAlaLys	967
DB	535	AGGACATCTCTGTTGGAGACCGAGACCCAGTCATAGATGCGTACTTCGGGCTGGCCAAAG	594
QY	968	LeuLeuProLeuGlnLysAspTyrTyrValValArgGluProGlyLysLeuProIlePhe	987
DB	595	CTGGTGCACAGACAAAGACTACTACGTCGTCACAGAGACCGGGGACAGCCCGGCTTC	654
QY	988	TyrTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSer	1007
DB	655	TGGTACACACCCAGTGCAGTGGCTGACCAACGGTCTTCTCCGAGCAATCCAGCTGTGGAGC	714
QY	1008	PheGlyValValLeuTyrGlnLeuPheThr	1017
DB	715	TTCGGGGTCTGCTCTATAGCTTTCACG	745

LOCUS BE916148 545 bp mRNA linear EST 29-SEP-2000  
 DEFINITION 60166673F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3966903 5',  
 mRNA sequence.  
 ACCESSION BE916148  
 VERSION BE916148.1 GI:10416498  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 545)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM9140 row: m column: 16  
 High quality sequence stop: 545.  
 Location/Qualifiers  
 1..545  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_image="3966903"  
 /clone\_lib="NCI\_CGAP\_Mam1"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="10 months, virgin"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: Salt;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"  
 BASE COUNT 93 a 176 c 166 g 110 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6.25e-79 Length: 545  
 Score: 879.00 Matches: 175  
 Percent Similarity: 97.25% Conservative: 2  
 Best Local Similarity: 96.15% Mismatches: 3  
 Query Match: 15.00% Indels: 3  
 DB: 12 Gaps: 0  
 US-09-397-967-16 (1-1099) x BE916148 (1-545)  
 Oy 307 ProLaGlgUgUHSArGleuValThrValPhrArgMetAspGlyHISleLeuGluAla 326  
 |||||||  
 Db 3 CCGGAGGAGGACCGCGCTGGTCTACTGTCAACGATGACGCGCCACATCTGGAAAGC 62  
 |||||||  
 Oy 327 GluPheProGlyLeuProGluAlaLeuSerPheValAlaLeu-ValAspGlyTyrPheAr 346  
 |||||||  
 Db 63 GAGTTTCGGGGGCTCGCTGAGGCGCTGCTTCTGCGCCCTCGGTGATGGTACTTCG 122  
 |||||||  
 Oy 346 GluLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProPheArgLeuGlu 366  
 |||||||  
 Db 123 CCGTATCTCGAGACTCCAGGCAATATTCTGCAAGAGAGTGGCGCGCCAGCGCTGTGA 182  
 |||||||  
 Oy 366 uGluGluAlaAspValCysHis-GlyProIleThrLeuAspPheAlaIleHisLysLeuL 386  
 |||||||  
 Db 183 GGAAGAGGCGAGCTGTGACAGTGGACCATCAGCTTGAAGCTTGCATCCACAAGCTGA 242  
 |||||||  
 Oy 386 ySaLaalAagLysSerLeuProGlyTyrThrTyrIleLeuArgSerProGlnAspTyrAsp 406  
 |||||||  
 Db 243 AGGCGGCTGGCTCCCTCCAGGACATATATTCCTCCGCCAGCCGACAGAGATATGACA 302  
 |||||||  
 Oy 406 eRpHeLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyLysLeuI 426

|||||  
 Db 303 GCTTCTCTTACCGCTGGCTCCAGACTCTCTTGGCCCCGACATCAAGAGGCTGCCTCA 362  
 |||||||  
 Oy 426 leArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerL 446  
 |||||||  
 Db 363 TCCGCGAGACCCCGAGGGGGCTTCTCCCTGGTTGGCCCTCACCGCCCAAGAGGCC 422  
 |||||||  
 Oy 446 euArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeu 466  
 |||||||  
 Db 423 TCGGGAGCTGCTTGCAGGCTGCTGGAATTCGTGGCTGCAGAGTACAGGTGCTGCCTGA 482  
 |||||||  
 Oy 466 yrluThrSerCysAlaAlaProArgProLysGluLysSerAsnLeuIleValAlaArg 486  
 |||||||  
 Db 483 ACTTAACATCTCTGCTGCTCCAGA-CCCAAGGAAAAAGTCAATTGATGCTGTGCGAA 541  
 |||||||  
 Oy 486 rg 486  
 ||  
 Db 542 GG 543  
 ||  
 RESULT 15  
 BE913247 539 bp mRNA linear EST 29-SEP-2000  
 LOCUS BE913247  
 DEFINITION 601668246F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:3968586 5',  
 mRNA sequence.  
 ACCESSION BE913247  
 VERSION BE913247  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 539)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM9144 row: k column: 11  
 High quality sequence stop: 539.  
 Location/Qualifiers  
 1..539  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_image="3968386"  
 /clone\_lib="NCI\_CGAP\_Mam1"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="10 months, virgin"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: Salt;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"  
 BASE COUNT 92 a 160 c 149 g 138 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.98e-78 Length: 539  
 Score: 874.00 Matches: 174  
 Percent Similarity: 97.21% Conservative: 0  
 Best Local Similarity: 97.21% Mismatches: 5  
 Query Match: 14.91% Indels: 2  
 DB: 12 Gaps: 0  
 US-09-397-967-16 (1-1099) x BE913247 (1-539)

```
OY 4 ProSerGIuIuThrProLeuIleProGIuArgSerCysSerLeuSerSerSerGIuAla 23
    |||||||
Db 3 CCAAGTGGAGACACCTGTGATCCCTCAGAGCTTTCAGACCTCATCTCAGAGCA 62
    |||||||
OY 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGIuArgLeuSerPhe 43
    |||||||
Db 63 GGAGCCCTGCAATGCTCTCCCTCCCGGGGACCTGGGCTCCCGACGATAGTCATTC 122
    |||||||
OY 44 SerPheGIuAspTyrLeuAlaGIuAspLeuCysValArgAlaAlaLysAlaCysGlyIle 63
    |||||||
Db 123 TC-TTTGGGACACTTCTGCTGAGAGATTATGTGCGAGCTGCCAAGGCCGTGGCATC 181
    |||||||
OY 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGIuAspPheSerCysTyrPhePro 83
    |||||||
Db 182 CTGGCTGTATCATTCGCTTTTCGCTCTGGCCACTGAGGACTTCTCTGTGTTCCC 241
    |||||||
OY 84 ProSerHisIlePheCysIleGIuAspValAspPheGIuValLeuValTyrArgLeuArg 103
    |||||||
Db 242 CCAAGCCACATCTTCTGATAGAGAGCTGAGACTCAAGT-CTGGTCTACAGGCTACGC 300
    |||||||
OY 104 PheTyrPheProAspTyrPheGIuLeuGIuThrCysHisArgPheGIuLeuArgLysAsp 123
    |||||||
Db 301 TTTTATTTCCCTGACTGTGTTGGGCTGAGACATGTCACCGCTCTGGGCTGCCAAGAT 360
    |||||||
OY 124 LeuThrSerAlaIleLeuAspLeuHisValLeuGIuHisLeuPheAlaGIuHisArgSer 143
    |||||||
Db 361 TTGACCAgTGCATCTTGACTTACATGTTTAGAACATCTCTTGTGCTCAGCACCGCAGT 420
    |||||||
OY 144 AspLeuValSerGIuArgLeuProValGIuLeuSerMetLysGIuGIuArgLysIuPheLeu 163
    |||||||
Db 421 GACCTGGTAGTGGGCGCTCCCGTGGGCTTAGCATGAGAGACAGGAGAGTTCTTG 480
    |||||||
OY 164 SerLeuAlaValLeuAspLeuAlaGIuMetAlaArgGIuGIuAlaGIuArgProGly 182
    |||||||
Db 481 AGCCTGGCCTGCTGACTTGGCGCAGATGGCTGAGAGAGGCGCCAGCCCGCAGGA 537
    |||||||
```

Search completed: April 28, 2003, 16:55:17  
Job time : 3456.75 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 28, 2003, 13:05:49 ; Search time 585.936 Seconds  
(without alignments)  
4223.916 Million cell updates/sec

Title: US-09-397-967-16  
Sequence: 1 MAPSEPTPLIPONCSLSS.....RPAFATLSPOLDPLMRGRPG 1099

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framed\_p2n\_model -DEV=xlp  
-Q=/cgnr\_1/USPFO.spool/US09397967/rumat\_26042003\_182312\_10702/app-query.fasta\_1.1486  
-DB=Geneseq\_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967.ecgn.1.1418 -rumat\_26042003\_182312\_10702 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEDUARY -NEG\_SCORES=10 -MAIT -DSHLOCK=100 -LONGLONG  
-DEV.TIMEOUT=10 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5223	89.1	4016	AA111083	Mouse JAK3 protein
2	4418	75.4	3807	AAT30862	Protein tyrosine k
3	2629	44.9	3435	AAV61801	JAK2 protein encod
4	2621	44.7	4482	AAAX80971	Human JAK2 kinase
5	2621	44.7	4482	AAZ58947	Human JAK2 kinase
6	2621	44.7	5117	AAD24311	Human JAK2 kinase
7	2619.5	44.7	3629	AAQ85412	Human JAK2 kinase
8	2619.5	44.7	3629	AAQ66244	JAK2 polynucleotid
9	2619.5	44.7	3629	AAQ66245	Murine (Janus kina
10	2376.5	40.6	3473	AAQ25307	JAK2 encoding DNA.
11	2360	40.6	3495	AAD23680	Murine JAK2 tyrosi
12	1908.5	32.6	4191	ABK72331	DNA encoding lymph
13	1904	32.5	3429	AAQ85413	Human JAK1 kinase
14	1904	32.5	3429	AAQ66245	Human JAK1 kinase
15	1904	32.5	3540	AAQ26039	Human (Janus kinas
16	1904	32.5	3540	AAQ26039	JAK1 encoding DNA.
17	1904	32.5	3541	ABK84065	Human CDNA differe
18	1904	32.5	3541	ABN95881	Gene #2379 used to
19	1904	32.5	3541	ABK72301	Lymphoma associate
20	1904	32.5	3541	ABL61826	Colon adenocarcino
21	1842	31.4	3538	ABK72332	DNA encoding lymph
22	1817.5	31.0	4176	ABK84572	Human CDNA differe
23	1817.5	31.0	4176	AAD24312	Human TYK2 DNA. H
24	1815.5	31.0	3561	AAQ85414	Human TYK2 kinase
25	1815.5	31.0	3561	AAQ66246	Human TYK2 kinase
26	1815.5	31.0	3561	AAQ66246	Human TYK2 kinase
27	1813.5	30.9	4248	AAK52035	Human polynucleoti
28	1806.5	30.8	4218	AAH98351	Human E2F-derived
29	1806.5	30.8	4220	AAK5019	Human polynucleoti
30	1564	26.7	4294	AAK85150	DNA encoding novel
31	1010	17.2	795	AAK85150	Human Janus kinase
32	948.5	16.2	726	AAK10807	Human Janus kinase
33	857.5	14.6	778	AAK10809	Human Janus kinase
34	816.5	13.9	2000	AAK78097	Human immune/haema
35	746	12.7	2327	AAZ3632	Human breast tumou
36	713	12.2	2338	AAK18071	Lung cancer associ
37	680.5	11.6	5045	ABK02929	Drosophila melanog
38	630	10.8	1026	AAK85149	DNA encoding novel
39	592.5	10.1	811	AAK10806	Human Janus kinase
40	507	8.7	3222	AAK78095	Human immune/haema
41	497	8.5	3127	AAK80650	DNA encoding novel
42	483.5	8.3	975	ABK72276	Lymphoma associate
43	483.5	8.3	975	ABK72333	DNA encoding lymph
44	482	8.2	3417	ABL91665	Human polynucleoti
45	477.5	8.1	3713	AAQ47065	He truncated rece

## ALIGNMENTS

RESULT 1  
AA111083 standard; CDNA: 4016 BP.

AA111083; 08-APR-1996 (first entry)

Mouse JAK3 protein-tyrosine-kinase gene.

Mouse; Janus kinase; JAK3; protein-tyrosine-kinase; cytokine;  
signal transduction; 3Dc13; reverse transcription; PCR;  
polymerase chain reaction; primer; interleukin-3;  
granulocyte-macrophage colony stimulating factor; CDNA probe;  
granulocyte colony stimulating factor; Escherichia coli;  
haematopoietic cell; differentiation; leukaemia; therapy; ss.  
Mus musculus.

```

XX Key Location/Qualifiers
FH 7..3906
FT CDS /*tag= a
FT /product= JAK3 protein-tyrosine-kinase
FT /note= "Ec-2.7.1.112"
XX
XX PN MO9601838-A1.
XX
XX PD 25-JAN-1996.
XX
XX PE 28-JUN-1995; 95MO-US08354.
XX
XX PR 08-JUL-1994; 94US-0272368.
XX
XX (UTEM ) UNIV TEMPLE.
XX
XX PI Rane SG, Reddy PE;
XX
XX DR MPI; 1996-097584/10.
XX
XX DR P-PSDB; AAR88560.
XX
XX PT JAK3 protein tyrosine kinase and DNA encoding it - useful to induce
XX differentialiation of haematopoietic cells to treat or control
XX leukaemia
XX
XX PS Claim 4; Page 26-28; 49pp; English.
XX
XX CC The sequence encodes a novel protein-tyrosine-kinase of the Janus
XX kinase (JAK) family (involved in cytokine signal transmission),
XX designated JAK3. The sequence has been isolated from mouse 32Dc13
XX cells growing in interleukin-3 or granulocyte-macrophage colony
XX stimulating factor by reverse transcription-polymerase chain
XX reaction amplification, using primers J3 and J4 (AAT11084-T11085),
XX based on conserved sequences in JAK JH2 and JH1 domains. A 950-bp
XX cDNA probe is isolated, which is then used to screen a granulocyte
XX colony stimulating factor-treated 32Dc13 cDNA library in phage
XX lambda-gli1 in Escherichia coli to isolate the full-length gene.
XX The JAK3 protein is useful for stimulating the differentiation of
XX haematopoietic cells in therapy or control of leukaemia.
XX
XX SQ Sequence 4016 BP; 742 A; 1203 C; 1235 G; 836 T; 0 other:

Alignment Scores:
Pred. No.: 0 Length: 4016
Score: 5223.00 Matches: 1060
Percent Similarity: 80.59% Conservative: 3
Best Local Similarity: 80.36% Mismatches: 32
Query Match: 89.13% Indels: 235
DB: 17 Gaps: 6

US-09-397-967-16 (1-1099) x AAT11083 (1-4016)
QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerSerLeuSerSer 20
DB 7 ATGGCACTCTCAAGTGGAGACACCTCTGATCCCTCAGGCTCTTCCAGCTCTCATCC 66
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40
DB 67 TCGAGGCGAGAGAGCCCTGCATGTCTCTCTCTCCCGGGGAGCTGGGCTCCCGCAGCA 126
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuGlyValArgAlaAlaAlaAla 60
DB 127 TTGTCACTCTCTTTGGGAGCTACTGCTGAGAGATTATGTGCGAGCTGCCAAGGCC 186
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
DB 187 TGTGCGATCTCTGCTGTTCATCTCGCTTTCCTGCGCAGCTGAGCACTCTCTTGC 246
QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
DB 247 TGGTTTCCCAAGCCACATCTCTGCTAGAGAGAGCTGAGACTCAAGTCTTGCTAC 306

```

```

QY 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
DB 307 AGGCTACGCTTTTATTTCCCGACTGGTGTGGCTGGAGACATGTACCGCTTTGGGCTG 366
QY 121 ArgIysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
DB 367 CGCAAGATTTGACCACTGACATCTTACTTACATGTTTATAGAACATCTCTTGTCTCAG 426
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
DB 427 CACCGCAGTACCTGGTGTAGTGGGCGGCTCCCGGTGGGCTTACCATGAAAGACAGGGA 486
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
DB 487 GAGTTCCTGAGCGCTGGCGGCTGCTGACTGGCCAGATAGCTGTGAGACAGCCACGCGC 546
QY 181 ProGlyIleLeuLeuLysThrValSerTyrIleLysLeuProProSerLeuArgAsp 200
DB 547 CCAGAGAGAGCTGTGAAGACAGCTGAGTTACAAACCTGTCTGCGCCAGCCTCGCGCAT 606
QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValIleAla 220
DB 607 GTGATCCAGGGCCAGAACTTCGTGACACGACGCGCATCCGACAGACCTGTGCTTGGCG 666
QY 221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240
DB 667 CTGC--GCGTGTGTCTCCCTCCAGCGGACCGCTACGCGCTCAGCCCAAGTATAT--CTG 723
QY 241 AspLeuGluArgGluHisProAlaIleThrThrGluThrPheArgValGlyLeuProGly 260
DB 724 GACTGTGAGGCGGCTACATCCAGCGGCCACACGACCTTCCTGTGGGCTCCCGGGC 783
QY 261 AlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSer 280
DB 784 GCCCAGGAGGAGCGGGGCTTCTGCTGTGGCGGGGCAACAGGCATCTCTGAGAGCTCC 843
QY 281 AsnAsp----- 282
DB 844 GGGGACAGAGAGTGTGCTGGGCTGGCTTGAGAACGGGGTCCGGGGAGCGCTGGCGG 903
QY 282 ----- 282
DB 904 TTACTAAAGAACCCGCTCCGACAAATCAGAAAGTGGCTAGTCCCGGGCGGCTGATTGC 963
QY 282 ----- 282
DB 964 ACTAGGGGGGAGGGGGGCGGGTCAATTAAAGTGGGTCCGGGAGTAGAGCTCCCGCAG 1023
QY 282 ----- 282
DB 1024 GGACTGACCTGGGGCTGTACTAGAAAGGGTGAGACTAGTAGGGGGCGGGACGACAGAA 1083
QY 282 ----- 282
DB 1084 GCGGATGTATGATTGGATGATGGGATGACCCGACAGGCTGTCTTCAAGTCCCTTA 1143
QY 282 ----- 282
DB 1144 ACCTGTGCCCCCCCCCACAAGTGGAGGTGGAGAGAGAGAAAGAGAGAGCTGTGCT 1203
QY 282 ----- 282
DB 1204 GGGTCAAGGCTGTAAATCCCTGTGCTTCTCTCTTTTCTTTTAAATGAT 1263
QY 282 ----- 282
DB 1264 TTTCGAGACAGGCTTTCTGTATAGCTGTGGCTGTCTGAACTCACTTTGTAAGACA 1323
QY 282 ----- 282
DB 1324 GCGTGGCCTGGAATCAGAAATCCGCTGCTGCTGAGATGCTGATTAAAGG 1383
QY 283 -----GluLeuPheGlnThrPheCys 289

```

```

Db 1384 CGTGTGCCACTACGCCGCTACTGCTGTCTCTTTCTTCTCAGCTTTTCCAGACTTCTGT 1443
Qy 290 AaPheProGluLeuValAspValSerTLeuSGLnaLProArGValGlyProAlaGly 309
Db 1444 GACTTTCCGAAATCGGTGATGTCACATCAACGACG--CCACGTGTGGTCCGAGAGG 1501
Qy 310 GluHISArgLeuValThrValThrArgMetAspGlyHISLeuGluValGluPhePro 329
Db 1502 AG-CACCGCTGTACTGTGACACAGGATGACGCGCACATCTCGAAGCGAGTTCCG 1560
Qy 330 GlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyThrPheArgLeuIleCys 349
Db 1561 GGGCTGCTGAGCGGCTGTCTTTCGTGGCCCTGTGATGGTACTTCCGCTGATCTGC 1620
Qy 350 AspSerArgHISThrPheCysIleGluValAlaProProArgLeuLeuGluGluAla 369
Db 1621 GACTCCAGGCAATTATTCTTGCAAGAGGTGGCGGCCACGCTGTGAGAGGAGGCG 1680
Qy 370 AspValCysHISGlyProIleThr----- 377
Db 1681 GAGCTGTGCATGACCCATCAAGTAAAGAGTATGGGGTCAACAGTGGGATGGGA 1740
Qy 378 -----LeuAspPhe 380
Db 1741 TGGGCAAGAGACTGTCTGGGTGGACCTGCTGTCTTGGCTTACGTTAGACTTT 1800
Qy 381 AlaIleHISLeuLeuValAlaAlaGlySerLeuProGlyThrThrIleLeuArgArgSer 400
Db 1801 GGCATCCACAGCTGAAAGGCCCTGCTGCTCCAGGACCTATATTCTCCGCGCAC 1860
Qy 401 ProGlnAspTrpArgSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420
Db 1861 CGGAGAGTATGACAGCTTCTTCTTACCGCTGCTGACAGTCCCTGTGGCCCGGAC 1920
Qy 421 TyrIleGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValIleLeuSer 440
Db 1921 TACAAAGGCTGCTCTATCCGCGAGGACCCAGCGGGCTTCTCCCTGGTTGC-CTCAC 1979
Qy 441 GlnProHISArgSerLeuArgGlyLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgVal 460
Db 1980 CAGCCCCACAGAAAGCTCGCGGAGCTGTGCAGC-TGTGGAATTC-GGGCTGGAGAG 2037
Qy 461 AspGlyAlaAlaLeuThrLeuThrSerCysAlaProArgProGlySulSerAsn 480
Db 2038 GAGGCTGCTCCCTAACAATCAATCTGCTCCAGACCCAGAAAGAAATCCAAAT 2097
Qy 481 LeuIleValAlaArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500
Db 2098 TTGATCGTGTGTCGAAGGGCTGCACCCCGGCTGTGCTGCTCCCGCTCTGCTGC 2157
Qy 501 CysAlaLeuThrGlnLeuSerPheHISThrIleProThrAspSerLeuGluTrpHISGlu 520
Db 2158 TGTGGCTGACACACACTGAGCTTCCACAAATTCACACAGGACGAGCTG---GACACAG 2214
Qy 521 AsnLeuGlyHISGlySerPheThrIlePheArgGlySerArgArgGluValAlaAsp 540
Db 2215 AACCTGGGTACAGCTTCTTTTACCAAGATCTTCGCTGCGCCGAGGCGGAGGTCGTGAT 2274
Qy 541 GlyIleThrHISAspSerGluValLeuLeuGlyValMetAspSerArgHISArgAsnGly 560
Db 2275 GGTGAGACACATGACTCGGAAGTCTCTGAAGGTCACTGAGCTCAAGCATCCGAACTGC 2334
Qy 561 MetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTrpProHISLeuVal 580
Db 2335 ATGGAGTCTTTCTGGAAGCGGCAAGCTTGATGAGCCCAAGATATCTACCCGACTGTG 2394
Qy 581 LeuLeuHISGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValIleu 600
Db 2395 TTACTGCAAGCGGCTGTGATGGCTGGACACAGCATCATGGTCCAGAAATTTGTGATCA 2454
Qy 601 GlyAlaIleAspMetTrpLeuArgIleGlyAlaGlyHISLeuValSerAlaSerTrpLysLeu 620
|||||
Db 2455 GAGACATTTGACTGTACTGCGCAAGCGTGGCCACCTGGGTGACCCAGCTGGAAGTGT 2514
Qy 621 GlnValThrLysGlnLeuAlaIleValAlaLeuAsnTrpLeuGluAspLysGlyLeuProHIS 640
Db 2515 CAGGTGACCAAGAGCTGCGCATATGCTTTAACTACTTGGAGACAAAGCCCTTCTTCA 2574
Qy 641 GlyAsnValSerAlaArgLysValLeuLeuAlaIleArgGlyGlyAspLysAsnProPro 660
Db 2575 GGCACAGTCTCAAGCAGGAGGAGTCTCCGCTCGGTGAGAGGGGGGTGATGGAATCCACT 2634
Qy 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680
Db 2635 TTCATTAACTAGTATCTGTGTGATGCTCCACAGCTGTGCTGAGCTGGAATGCTCAC 2694
Qy 681 AspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700
Db 2695 GACAGAAATACCTGGGTGGCGCCCAAGTCTCCAGAGAGGCTCAACACTGTCTGGAG 2754
Qy 701 AlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHIS 720
Db 2755 GCTGACAGTGGGCTTTGGAGCCACACAGTGGAGGTTCACCGGGGACCCCGCAC 2814
Qy 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTrpGluAspGlnGlyGlnLeuPro 740
Db 2815 ATCACCCTGCTGAGCCCGCCCAAAAGCTGAAGTCTTATGAGACACAGGACAGCTGCC 2874
Qy 741 AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTrpAspProGly 760
Db 2875 GCTTCAAATGACAGAACTGCGGGACTTATCACACAGTGCATGGCTATGATCATCTGCC 2934
Qy 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTrp 780
Db 2935 CGGCGCCCTCTCTTCCGAGATATCTTCAAGACCTCAACAGGCTCATTAATCAATTAAC 2994
Qy 781 GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800
Db 2995 GAGCTCTCTCCA-GACCCACACCTGCTGCAATCCGAGTCTCTCGAGATGACTGTGGCT 3049
Qy 801 GlyAlaGlnLeuThrAlaCysGlnAspProAlaIlePheGluGluArgHISLeuLysTrp 820
Db 3050 GCGCGCCAGCTCTATGCTGTGCAGAGACCCCGCCATATC-GAGGAGAGACACTTAAGTAC 3108
Qy 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTrpAspProLeu 840
Db 3109 ATCTTTGTGGGCAAGGCACTTTGGACGCGTGGAGCTGTGCTGCTATGACCCCTGT 3168
Qy 841 GlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHISSerValProAspGln 860
Db 3169 ---GACAAATACGGGACCCCTGTGTGCAATGAAACAGTACAGCACAGCGGCGCAGACCAG 3225
Qy 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHISSerAspPheIleVal 880
Db 3226 CAGAGGGACTTCCAGCGGGAGATTCAGATCTTAAAGCTCTGCAACAGGACTTATGCTGC 3285
Qy 881 LysTrpArgGlyValSerTrpGlyProGlyArgGlnSerLeuArgLeuValMetGluTrp 900
Db 3286 AAGTACCGGGGAGTACAGCTATGCGGCCAGAGTCCAGAGCTGCGGTTGGTATGAGATAC 3345
Qy 901 LeuProSerGlyCysLeuArgAspLeuLeuGlnArgHISArg---GlyLeuHISThrAsp 919
Db 3346 CTGCGCCAGCGGCTCTGCGAGACTTCTTCCAGAGGCGCATGCGGGCGCTGCAACCGAC 3405
Qy 920 ArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlnMetGluTrpLeuGlyAlaArgArg 939
Db 3406 CGCTTACTGTGCTTCGCTTGGCAGATGTCGAAGGCAATGGAGTACCTGGGCTCGCGCCG 3465
Qy 940 CysValHISArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHISValLys 959
Db 3466 TGCCTACACCTGACCTGGCTGGCGCCCAACATCTTGTGGAGAGCGAGCTCATATGAAG 3525
Qy 960 IleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTrpTrpValValArg 979
Db 3526 ATGCGGAGCTTGGGCTCGCTAAGCTGTGCTGCGGAAAGGACTATCACTAGTGTGCTGC 3585
```

QY 980 GluProGlyInSerProIlePheTrpTyrAlaProGluInSerLeuSerAspAsnIlePhe 999  
|||||  
DB 3586 GAGCTGGCCAAAGCCCATCTTTGGTATGCCCCGGAGTCCCATCTGACAAACATCTTC 3645  
QY 1000 SerArgInSerAspValTTPSerPheGlyValValLeuTyrGluLeuPheThrTyrCys 1019  
|||||  
DB 3646 TCCCGCCAACTTGACGTGTGAGGCTCGAGGTGTGTGACGAGCTTTCACCTACTGCG 3705  
QY 1020 AsplSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGly 1039  
|||||  
DB 3706 GACAGAGCTCAGCCATCCATCGGTAGTCTCTGCGCATGATGGGGCTGAGCCGTAAGA 3765  
QY 1040 ProProLeuGlySerArgLeuLeuGluLeuLeuAlaGluGlyArgGlyLeuProProPro 1059  
|||||  
DB 3766 CCCCCGCTCTCCCTCTCTGAGCTGTGAGAGGGCCGACGCTCCACACCTCTCC 3825  
QY 1060 ThrCysProThrGluValGluGluLeuMetGluLeuGlySerTPAlaProGluProHisAsp 1079  
|||||  
DB 3826 ACCTGACCCACGAGGTTTCAAGAGCTCATGCACTGTGCGGGGCCAG-CGCGACGAC 3884  
QY 1080 ArgProAlaPheAlaThrLeuSerProGluInLeuAspProLeuTTPArgGlyArgPro 1098  
|||||  
DB 3885 CGGCGAGCTTCGCGACCTGAGCCGCCAGCTGAGAGGCTGTGCGTGAAGACCC 3941  
RESULT 2  
AAT30862  
ID AAT30862 standard; DNA; 3807 BP.  
AC AAT30862;  
XX 13-SEP-1996 (first entry)  
XX protein tyrosine kinase JAK3 CDNA.  
DE  
XX  
XX JAK3; protein tyrosine kinase; cell proliferation; differentiation;  
KW signal transduction; leukaemia; aplastic anaemia; myelodysplasia;  
KW polythemia vera; thrombocytosis; gene therapy; diagnosis; ss.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 168..3359  
FT CDS /\*tag= a  
FT polyA\_signal .3692..3696  
FT /\*tag= b  
FT polyA\_signal 3793..3797  
FT /\*tag= c  
XX  
XX MO9618639-A1.  
XX  
XX PD 20-JUN-1996.  
XX  
XX PF 15-DEC-1995; 95WO-US16435.  
XX  
XX PR 15-DEC-1994; 94US-0357598.  
XX  
XX PA (UWJO ) UNIV JOHNS HOPKINS SCHOOL MED.  
XX  
XX PI Clivin CI, Safford MG, Small D;  
XX  
XX DR WPI; 1996-300568/30.  
XX  
XX P-PSDB; AAR96037.  
XX  
XX PT Protein tyrosine kinase, JAK3, protein and nucleic acid - used in  
XX the gene therapy of cellular proliferative diseases, e.g. leukaemia,  
XX aplastic anaemia etc.  
XX  
XX PS Claim 4; Page 43-47; 97PP; English.  
XX  
CC A cDNA clone (AAT30862) codes for JAK3 (AAR96037), a new member of the  
CC JAK family of non-receptor protein tyrosine kinases, that probably  
CC plays a role in growth factor modulated differentiation.

CC proliferation and survival of haematopoietic stem/progenitor cells.  
CC It was derived from CD34+ mRNA obd. from normal human bone marrow  
CC by PCR amplification using primers (see also AAT30863-64) based  
CC on highly conserved motifs from protein tyrosine kinase catalytic  
CC domains. The JAK3 coding sequence can be incorporated into a  
CC vector and used for prodn. of recombinant JAK3. It can also be  
CC used in gene therapy protocols for leukaemia, myelodysplasia,  
CC polythemia vera, thrombocytosis and aplastic anaemia, or to  
CC stimulate haematopoietic cell proliferation. The gene was  
CC localised to chromosome 19, band p12-13.1.  
XX  
XX SQ Sequence 3807 BP; 716 A; 1219 C; 1099 G; 773 T; 0 other;  
SQ  
Alignment Scores:  
Pred. No.: 0 Length: 3807  
Score: 4418.00 Matches: 870  
Percent Similarity: 86.948 Conservative: 62  
Best Local Similarity: 81.168 Mismatches: 130  
Query Match: 75.398 Indels: 12  
DB: 17 Gaps: 6  
US-09-397-967-16 (1-1099) x AAT30862 (1-3807)  
QY 1 MetaLabProProSerGluGluThrProLeuIleProGluInArgSerCysLeuSerSer 20  
DB 168 ATGACACCTCCAAAGTGAAGAGAGCCCTGATCCCTCAGGTTTATGACACCTTTGTC 227  
QY 21 SerGluInGluAlaLeuHisValLeuLeuProProArgGlyProGlyProProGluArg 40  
DB 228 ACGAGGCTGTGCTCCCTGCTCATGCTGCTGCTCCGCTGGGGGCCCGGCCCAAGCGC 287  
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuGlyValArgAlaAlaValAla 60  
DB 288 CTATCTTCTCTCTTGGGGGACCACTTGCTGAGAGACCTGTGCGGCTGCCAAGGCC 347  
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuArgThrGluAspPheSerCys 80  
DB 348 AGCGGATCTGCTGCTGTGATCCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 407  
QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThr-GluValLeuValTyr 100  
DB 408 TGGTT-CCCCGAGACCAATCTTCCGTGAGAGATGCCAGACCCCAACATCTCTGTGTA 466  
QY 100 ArgLeuArgPheTyrPheProAspTTPPheGlyLeuGluThrCysHisArgPheGlyLe 120  
DB 467 CAGATTCGCTTTACTTCCCAATGTGTGGCTGAGAGAGTCCACGCTTGGGCT 526  
QY 120 uArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaG 140  
DB 527 ACGCAAGATTTGGCCAGATGCTATCTTGACCTGACGCTCGAGCACCTTTGCCCA 586  
QY 140 nHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnI 160  
DB 587 GCACGCGAGTACCTGAGTGGAGGCGGCTCCCGCTGCTCAGTCTCAAGAGGAGGAGG 646  
QY 160 yGluPheLeuSerLeuAlaValLeuAspLeuAlaGluMetAlaArgGluGlnAlaGlnAr 180  
DB 647 TGAGTCTCTCAAGCTGCGCGGTGTGAGCTGCGCGGATGCGGAGAGCAGGCGGAGGCG 706  
QY 180 gProGlyGluLeuLeuTyrThrValSerTyrLysAlaCysLeuProProSerLeuArgAs 200  
DB 707 GCGGGAGAGAGCTCGTGAAGAGCTGACCTACAGAGCTGCTACCCCAACCTGCGCGGA 766  
QY 200 pValIleGlnGlyGlnAspPheValThrArgArgGlyLeuArgThrValValLeuAl 220  
DB 767 CCGATTCACAGGCGCTGAGCTTGTGACGCGGAGGCGATTCGAGAGAGGCTGAGAGCCC 826  
QY 220 aleuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleL 240  
DB 827 CCGTGC-CGCGGTGCGCGCTGCGAGGAGAGCGGCACTGCTCATGCGCAAGTACATCA 885  
QY 240 euAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260  
:::|||||



Db 886 TGGACCTGGAGCGCTGGATCCAGCCGGGGCCGAGACCTTCACGTGGCCCTCCG 945  
Qy 260 IYALA-----GIngluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleP 277  
Db 946 GGGCCCTGGTGGCCACGACGGGGCTGGGGCTGCTCCGGCTGGTGGTACGGGGCATG 1005  
Qy 277 rOfPserSerAsnsp---GluLeuPheGlnThrPheCysAspPheProGluIleValA 296  
Db 1006 CCTGGACCCGAGGAGAACGAGAGCTCCAGCCCTTCGCGACTTCACAAATGTAG 1065  
Qy 296 sPValSerIleAsnGlnAlaProArgValGlyProIleGlyIleHisAspGluValThyV 316  
Db 1066 ACATTAAGCTCAACGACGGCCCGCGGTGGCCGGCCGAGACACCGCCTGGTACG 1125  
Qy 316 aLThrArmeLaspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeus 336  
Db 1126 TTACACGAGACACAAACAGATTTTAGAGCGGAGTTCCAGGGCGTGGCCGAGGCTCT 1185  
Qy 336 eRPhValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisIlePheC 356  
Db 1186 CATTGGCTGGCTGGAGCGGCTACTTCGGCTGACACGAGACTCCAGACTTCCTCT 1245  
Qy 356 yALysGluValAlaProProArgLeuLeuGluGluGluAlaAspValCysHisGlyProI 376  
Db 1246 GCAAGGAGGTG---GACCCGAGGCTGCTGAGAGAGTGGCCGAGCATGCCACGGCCCA 1302  
Qy 376 LeThrLeuAspPheAlaIleHisIleCysLeuLysAlaAlaGlySerLeuProGlyThyTyrI 396  
Db 1303 TCACCTCTGACCTTGGCATCAACAGCTCAAGACCTGGGGGCTCAGCTCGTGGCTCTTAG 1362  
Qy 396 LeLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrP 416  
Db 1363 TTCTCGCGCGCATCCCGCGAGCTTGGACAGCTTCCTCCTCACTGCTGTGCCAAMAC 1422  
Qy 416 rOLeuGluProAspTyrLysGlyLysLeuIleArgGlnAspProSerGlyAlaPheSerL 436  
Db 1423 CCTTGGCTCTATTTAAAGGCTGCTCATCGGGCGACGCCACACAGGAACCTTCTTC 1482  
Qy 436 euValGlyLeuSerGlnProHisIleArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsn 456  
Db 1483 TGGTGGCTCAAGCCGACCCACAGACAGTCTTGAAGCTCCGAGCACTGGAGCTGGAG 1542  
Qy 456 eRGlYLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaAlaProArgProI 476  
Db 1543 GGGGGCTGCAGCTAGATGGGGTGGCAGTACCCTCCTCTGCTATGCCAGACCA 1602  
Qy 476 yAGLILySerSerAsnLeuIleValAlaArgArgGlyCysAsnProAlaProAlaProGlyC 496  
Db 1603 AAGMAAGTCAACCTGATTTGGTCCAGAGAGTACAGCCACCCACATCATCTTGG 1662  
Qy 496 yASerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisIleProThrAspS 515  
Db 1663 TTCAGGCCCAATCCCAATACAGCTGAGTACAGATTTTCAACAGATCCCGTGCAGCA 1722  
Qy 515 eRLeuGluThrPheIleGlnAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySera 535  
Db 1723 GCCTGAGTGGATGAGAACCTGGGGCATGGGCTCTTCCACCAAGATTTACCGGGCTGTC 1782  
Qy 535 rGArgGluValAlaAspGlyLysIleThrHisAspSerGluValIleLeuLysValMetAsp 555  
Db 1783 GCCATAGAGTGGTGGAGGGAGGCCGAAGACAGAGTGGCTGGCAAGGTATGAGG 1842  
Qy 555 eRArgHisIleArgAsnCysMeGluSerPheLeuGluAlaIleSerLeuMetSerGlnValS 575  
Db 1843 CCAAGCACAAGAACTCATGAGTCTTCTCGAAGCAGACAGCTTGAGAGCAAGTGT 1902  
Qy 575 eRTyrProHisIleuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValG 595  
Db 1903 CATTACCGCATCTGCTGCTGCACAGGGCTGATGCTGGAGACACACATGATGTCG 1962  
Qy 595 InGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisIleValS 615  
Db 1963 AGGATTTGTACACCTGGGGGCTATGACATGATCTGCAAAACCTGGCCACACTGGTGC 2022

Qy 615 eRAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluA 635  
Db 2023 CAGCCAGCTGGAGAGCTGACAGTGTGTCAAAACAGCTGGCCCTACGCCCTCACTATCTGAG 2082  
Qy 635 sPLysGlyLeuProHisIleGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyC 655  
Db 2083 ACAAGGCGCTGTCCCATGTGCAAATGTCTGCCCCGGAAGGTGCTCGGCTCGGAGAGGG 2142  
Qy 655 LYAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeus 675  
Db 2143 CTGATGGAGCCCGCCCTTATCAAGCTGAGTACCCTGGGGTCACCCCGCTGTATTA 2202  
Qy 675 eRLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaG 695  
Db 2203 GCTGAGATGTGCACACGACAGATCCCTGGGTGGCCCGCCGAGTGTCTCCGGAGAGCCG 2262  
Qy 695 InThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheC 715  
Db 2263 AGCACTTAGCTTGGAACTGACAAAGTGGGGCTTCGGCCACAGGCTGTGGAGTGTTA 2322  
Qy 715 InArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluA 735  
Db 2323 GTGGCTCACCATGCCATCAGTGGCCCTGAGTCTGTCAAGAAATTCCAATTTTATGAG 2382  
Qy 735 sPGLInGlyLInLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysM 755  
Db 2383 ACCGGCAGACAGCTGTGGCCGCCCAAGTGGACAGAGCTGGCCCTGCGATTCACACATGCA 2442  
Qy 755 eRAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyL 775  
Db 2443 TGGCCATATAGCGCGGTGACAGAGCCCTCTTACGACCCCTCATTCGTGACCTCAATAAG 2502  
Qy 775 euIleThrSerAspTyrGlyLeuLeuSerAspProThrProGlyIleProSerProArg 795  
Db 2503 TCATCTCTTCAAGCTATGAGCTCTCTCAAGACACAC---TGGTCCCTGGCAGCTGTCG 2559  
Qy 795 sPGLInCysValAlaGlyAlaGlnLeuTyrLacCysGlnAspProAlaIlePheGluG 815  
Db 2560 ATGGGCTGTGG---AATGGGCCACACTGTATGCCGCAAGACCCACAGATCTTGAG 2616  
Qy 815 InArgHisIleuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeu 835  
Db 2617 AGAGACACCTCAAGTATCATCTCACAGCTGGGCAAGGGCTTCTTGGCACAGTGGACCTGT 2676  
Qy 835 yARgtTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln 855  
Db 2677 GCCGCTATGACCCGCTAGCGGACAAATACAGTGCCTGGTGGCCGTGAACAGCTCAGC 2736  
Qy 855 IASerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeu 875  
Db 2737 ACAGCGGGCCAGACACAGAGAGGACTTTCAGGGGAGATTCAGATTCCTCAAAAGCAGAG 2796  
Qy 875 IASerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuA 895  
Db 2797 ACAGTATTTCAATTCAGATATCTGTGTGCTGAGCTATGGCCCGGCCGACAGAGCCCTG 2856  
Qy 895 rGLeuValMetGluTyrTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg 915  
Db 2857 CGCTGCTATGAGATACCTGCCCCAGCGGCTGTGGCGACTTCTTGCAGCGGCCACCGGG 2916  
Qy 915 LYLeuHisIleThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGlnTyrL 935  
Db 2917 GCCTGATGCGACAGCGCCCTCTTCTATTTCCCGAGAGATTCGCAAGGGCATGAGTACC 2976  
Qy 935 euGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerC 955  
Db 2977 TGGGCTCCCGCGCTCGTGCACCGGACCTGGCGCCGCAAAACATCTCTGTGAAGAGG 3036  
Qy 955 InAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspT 975  
Db 3037 AGGCACAGCTCAAGATGCTGACTTGGCTTGAAGTGTGCGGCTTGCACAAAGACT 3096

QY 975 YTYValValaIValArgLupProGlyInserProIlePheTrpTyrAlaProGlyLeuS 995  
 |||  
 Db 3097 ACTACGIGTGTCCGGAGAGCCAGAGCCCATTTTGTGTATGCCCCCAATCCCT 3156  
 QY 995 erAspaSnIlePheSerArgInSerAspValITrpSerPheGlyValValLeuTyrGluL 1015  
 |||  
 Db 3157 CGGACACATCTCTCTCGCAGTACAGCTGTGAGCTTGGGGTCTCTCTGACGAGC 3216  
 QY 1015 eupheTrpTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetGlyP 1035  
 |||  
 Db 3217 TCTTACCTACTGTGACAAAGGTGACGCCCTGCGGAGTTCTCTGCGATGATGGAT 3276  
 QY 1035 roGluArgGluGlyProProLeuGlyCysArgLeuLeuGluLeuAlaGluValArgTGL 1055  
 |||  
 Db 3277 GTGAGCGGAGATGCCCGGCTGTGCGGCTCTTGGAATCTGTGAGSAGSGCCAGAGC 3336  
 QY 1055 eupProProProTrpCysProThrGluVal 1065  
 |||  
 Db 3337 TGCGGGCGCTCTGCTGCTGCTGCTGAGTG 3368  
 RESULT 3  
 AAV61801  
 ID AAV61801 standard; cDNA: 3435 BP.  
 XX  
 AC AAV61801:  
 XX  
 DT 20-JAN-1999 (first entry)  
 XX  
 DE JAK2 protein encoding cDNA.  
 XX  
 XX JAK2 protein: cytoplasmic domain; beta subunit; screening; asthma;  
 KW Interleukin; granulocyte macrophage-colony stimulating factor; GM-CSF;  
 KM IL-3; IL-5; human; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 10..3429  
 FT /\*tag= a  
 FT /product= "JAK2 protein"  
 XX  
 PN WC9843087-A1.  
 PD 01-OCT-1998.  
 XX  
 PE 23-MAR-1998: 98WC-US05387.  
 XX  
 PR 24-MAR-1997: 97US-0041511.  
 XX  
 PA (PHAA ) PHARMACIA & UPJOHN CO.  
 XX  
 PI Kayles PS, Roberds SL;  
 XX  
 DR WPI: 1998-532151/45.  
 DR P-PSDB: AAW/6425.  
 XX  
 PT Screening for compounds useful for preventing or treating asthma -  
 PT by determining if compounds inhibit binding of the JAK2 protein to  
 PT e.g. IL-3, IL-5 or GM-CSF  
 XX  
 PS Disclosure: Fig 22A-C; 112pp; English.  
 XX  
 CC This cDNA encodes the full-length JAK2 protein. The N-terminal fragment  
 CC of JAK2 protein can be used in the method of the invention of screening  
 CC for compounds useful for treating or preventing asthma. The method  
 CC comprises contacting a molecule comprising at least the N-terminal 294  
 CC amino acid residues of the JAK2 protein, with another molecule comprising  
 CC at least 13 membrane-proximal cytoplasmic amino acids of interleukin  
 CC (IL)-3, IL-5 or granulocyte macrophage-colony stimulating factor (GM-CSF)  
 CC proteins in the presence of the candidate compound, and determining  
 CC whether the first and the second molecules form a complex. If the  
 CC compound inhibits complex formation, it can be used to treat asthma.

SQ Sequence 3435 BP; 1157 A; 598 C; 733 G; 947 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 4.05e-191 Length: 3435  
 Score: 2629.00 Matches: 534  
 Percent Similarity: 66.21% Conservative: 193  
 Best Local Similarity: 48.63% Mismatches: 340  
 Query Match: 44.86% Indels: 31  
 DB: 19 Gaps: 14  
 US-09-397-967-16 (1-1099) x AAV61801 (1-3435)  
 QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGly 39  
 |||  
 Db 111 TCCAGTTCTTACGAT-----GTATCTTACCATTCCCTTGGGAATGTGAGGCGA 161  
 QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAla 59  
 |||  
 Db 162 TTATCTGACCTTTCACATCTGCGGAGATGTGCGAGAAATCTGTATCTGCTCTTAA 221  
 QY 59 sAlaCysGlyLeuLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPhe 79  
 |||  
 Db 222 AGCTTGAGTATCACACCTGTGTATCATATATGTTTCTTAAATGAGTGAACAGAAAG 281  
 QY 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAspThrGluValLeu 99  
 |||  
 Db 282 GATCTGATATCCACCCACCATGATCTTCCATATGATGATGATGATGATGATGATGAT 341  
 QY 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluTrpCysHis 117  
 |||  
 Db 342 CTACAGAAATAGATTTTATCTTCTCTGTTATGTCAGTGCAGACAGACAGACAGAC 401  
 QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHis 137  
 |||  
 Db 402 GCATGGAATATCTCGAGTGTGTAAGCTCCTCTTGTATGATCTTGTGATGATGATGAT 461  
 QY 137 uPheAlaGluHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMet 157  
 |||  
 Db 462 CTTCCTCAGTACGCGGCGATGATTTTGTGATGATGATGATGATGATGATGATGATG 521  
 QY 157 sGluGluGlyLeuPheLeuSerLeuAlaValLeuAspLeuAlaGluMetAlaArgGlu 177  
 |||  
 Db 522 AACACAGGAGAAATGTCTTGGATGCGAGTGTGATGATGATGATGATGATGATGATG 581  
 QY 177 nAlaGluArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProPro 197  
 |||  
 Db 582 CGATCAAAACCCACTGGCGCATCTATACCTATACAGTACAGACAGACATCTTACCA 641  
 QY 197 rLeuArgAspValIleGluGlyGluAsnPheValThrArgArgIleArgGlyThr 217  
 |||  
 Db 642 TATTCAGCAAAAGATCCAGACTATGATTTTGTGCAAGGAGGATTAAGTACAGATT 701  
 QY 217 lValLeuAlaLeuLeuProGlyArgLeuProGlyArgProTyrAlaLeuMetAla 237  
 |||  
 Db 702 TCGCAGATTTTATGACCATTCAGCCATGCAAGCCACTGCCAANAATTGAACTTAA 761  
 QY 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal 257  
 |||  
 Db 762 GTATCTTATTAATCTGGAACCTGTGAGTCTGCTTACACAGAGAAATTTGAAGTAA 821  
 QY 257 yLeuProGlyAla-----GluGluGluProGlyLeuLeuArgValAlaGlyAs 273  
 |||  
 Db 822 AGAACCTGGAGGTGCTCTTCAGGTGAGAGATTTTGTCAACCATTTATTAATACTG 881  
 QY 273 pAsnGlyIleProTrpPheSer-----AsnAspGluLe 284  
 |||  
 Db 882 CGGTGGAATTTAGTGTGTAAGAGGAAACATAAAGAAAGTACAGACCTGACAGACAGA 941  
 QY 284 uPheGluThrPheCysAspPheProGluIleValAspValSerIleAsnGluAlaPro 304  
 |||  
 Db 942 TTTACAGTTATATTGCGATTTTCTTAATATATTGATGATGATGATGATGATGATG 1001  
 QY 304 gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetLysGlyHisIle 324

Db 1002 AGAGCGT---TCAAATGAAGCCGAGTTGTACTATCCATAAGCAAGATGGTAAATCT 1058  
Qy 324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyr 344  
Db 1059 GGAATTCACCTTACCTTATTAAGGAGCTTGTCTTCCGTTCAATTAATTAATGATGATA 1118  
Qy 344 rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLe 364  
Db 1119 TTATGATATTAACTGCAGATGCACATCATTTACCTGTAAAGAGATGACCTCCAGCGCT 1178  
Qy 364 uLeuGluGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLys 384  
Db 1179 GCTTGAAATATACAAAGCAACCTGCATGCGCCCAATTCGATGAGATTTTGGCATAGTAA 1238  
Qy 384 sLeuValAlaAlaGlySerLeuProGlyThrTyrIleLeuArgAspSerProGluAspGly 404  
Db 1239 ACTGAGAAAGACGAGTATACAGACTGACCTGATGTACTGATCCAGTCCTAAGGAGCTT 1298  
Qy 404 rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCys 424  
Db 1299 TAATTAATATATTTTGTGACTTTTGTCTGCGAGGAGAAATGTCATTAATATAAACACG 1358  
Qy 424 sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArg 444  
Db 1359 TTTGATACAAAAATGAGAAATGAGAGATACAACTCAGTGGACAAAGAAAGAACTTCAG 1418  
Qy 444 gSerLeuArgGluLeuLeuAlaAlaCysStrProAspSerGlyLeuArgValAspGlyAlaAl 464  
Db 1419 CAGCTTTAAAGTCTTTTGAATTTGTACCAGATGGAACCTGCTCCAGCAATATATAT 1478  
Qy 464 aLeuTyrLeuThrSerCysAlaAlaProArgProLysGluLysSerAsnLeuIleValVa 484  
Db 1479 TTTCAGATTTACTAAATGCTGCCCCCAAGCAAGATTAATCAAACTTCTAGTCTT 1538  
Qy 484 LArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAlaLe 503  
Db 1539 CAGAACGAATGGTCTTTGTGATGTACCAACTCACCACATTTACAGAGCCCTACTCAT 1598  
Qy 503 uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrPheHisGluAsnLeu 523  
Db 1599 GACCAAAATGGTGTTCACAAATGAGAAATGAGATTTGATTTAATGAAGAGCCCTGG 1658  
Qy 523 yHisGlySerPheThrLysIlePheArgGlySerArgGluValAlaAsp---GlyGln 542  
Db 1659 CCAAGGCACTTTACAAAGATTTTAAAGCGCTACGAGAGAGATGAGAGACTAGCGTCA 1718  
Qy 542 uThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGln 562  
Db 1719 ACTGCATGAACAGAAAGTCTTTTAAAGTCTGATGAAGCACACAGAACTATTTCAGA 1778  
Qy 562 uSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLe 582  
Db 1779 GCTTTTCTTTGAGCAGCAGAGATGATGAGCAAGCTTTCTCACAAGCATTTGGTTTAA 1838  
Qy 582 uHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyrLeuGln 601  
Db 1839 TTATGCAATATGTCTGTGCGAGAGCAGAAATATTTCTGCTCGAGACTTTGTAAATTTGCG 1898  
Qy 601 yAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrPheLysLeuGln 621  
Db 1899 ATCACTAGATACATATCTGAAAGAAAGAAATTAATTAATATATATGGAAGAACTTGA 1958  
Qy 621 nValThrLysGlnLeuAlaIleTyrAlaLeuAsnTyrLeuGlnLysAspLysGlyLeuProHisGln 641  
Db 1959 AGTTGCTAAACAGTTGGCATGGCCATCATTTTCTAGAAGAAACACCCCTTATTCATGG 2018  
Qy 641 yAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyLys-----AspGlyAsnPro 659  
Db 2019 GAATGTATGTGCAAAATATTTCTGCTATATCAGACAGAGACAGAGACAGGAAATCC 2078  
Qy 659 oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLe 679  
|||||

Db 2079 TCCTTTCATCAAACTTAGTATCTGGCATTAGTATTAACAGTTTTCGCAAGACATTTCT 2138  
Qy 679 uThrAspArgIleProThrPheValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe 699  
Db 2139 TCAGAGAGAAATACCATGGGTATCCACCTGAAATGCAATGCAAAATTTAAATTT 2198  
Qy 699 uGluAlaAspLysTyrPheGlyValAlaThrThrTyrPheGluValPheGlnArgGlyProAl 719  
Db 2199 GGCACAGCAAAATGAGATTTTGGTACACCTTTGTGGGAAATCTGCAGTGGAGAGATTA 2258  
Qy 719 aHisIleThrSerLeuGluProAlaLysLeuLysLeuLysPheTyrGluAspGlnGlyGlnLe 739  
Db 2259 ACCCTAACTGCTCTGATTTCTCAAGAAAGCTTACAAATTTTATGAAATAGACATCAGCT 2318  
Qy 739 uProAlaLeuLysTyrPheGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759  
Db 2319 TCCTGCACCAAAAGTGGCAGAAATTTAGCAACCTTTTAATTAATTTGATGATTAAGACC 2378  
Qy 759 oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779  
Db 2379 AGATTTACAGGCTCTCTTCAGAGCCATCATACAGATCTTAACAGTTTGTTAACGCCAGA 2438  
Qy 779 rTyrGluLeuLeuSerAspProThrProGlyLysProSerProArg---AspGlyLeuGly 798  
Db 2439 TTATGACATTTTAACAGAA---AATGACATGTACCAAAATAGAGAGATGAGTGCCTAGC 2495  
Qy 798 sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLe 818  
Db 2496 GTTTTCTGGTCC-----TTTGAAAGACCGGAGATCTTACACAGTTTGAAGAGACATTT 2549  
Qy 818 uLysTyrIleSerLeuLeuGlyLysGlyAsnPheLysSerValGlyLeuCysArgTyrAs 838  
Db 2550 GAAATTTTACAGCAACTTGGCAGAGGATATTTTGGAGATGAGAGATGTCGCCGATGA 2609  
Qy 838 rProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPro 858  
Db 2610 CCTCTTACAGCAACACACTGGGAGGTGCTCCTGTAAAGCTTTCACCAATAGCTGA 2669  
Qy 858 oAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPro 878  
Db 2670 AGAGCACCTTAAGACATTTGAAAGGAAATTAATCTCGAAATCTCTACACATGACAA 2729  
Qy 878 eIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMe 898  
Db 2730 CATGTGAAGATCAAGAGGAGTGTGCTACAGTCTGTCGGCGGATTAATTAATTAAT 2789  
Qy 898 tGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHis 917  
Db 2790 GGAATATTTTACATATGGAAGTTTACAGACTATCTTCAAAACATTAAGAACGCGATGA 2849  
Qy 917 sThrAspArgLeuLeuLeuPheAlaIleProGlnIleCysLysGlyMetGlyTyrLeuGlyAl 937  
Db 2850 TCACATTAACACTTGTGACATACATCTCAGATATGCAAGGATATGAGATATCTTGGTAC 2909  
Qy 937 rArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGlnAlaHis 957  
Db 2910 AAAAGATATATTCACAGGAGATCTGCAACAGAAATATATGTGTGTGAGACAGAGAACAG 2969  
Qy 957 sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVal 977  
Db 2970 AGTTAAATTTGAGATTTTGGGTAAACCAACTCTTCCACAAAGCAAGAAATTAATTAAT 3029  
Qy 977 lValArgGluProGlyLysLeuSerProIlePheThrPyrAlaProGlyLysLeuSerAspAs 997  
Db 3030 ACTTAAGAGACCTGGTGAAGTCCCATATTTGTGTAGTGTCCAGATCTACAGACAGAGG 3089  
Qy 997 nIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluIlePheThr 1017  
Db 3090 CAAGTTTCTGTGGCTCAGATCTTGGAGCTTTGGAGCGTTCGTATGAACCTTTTCC 3149  
Qy 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArg 1037  
Db 3150 ATACATTTGAGAGATTAAGTCAACAGCGGAATTTATGCTATGATGATGCAATGACAA 3209  
|||||

QY 1037 gGluGlyProProLeuGly---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuP 1056  
 Db 3210 ACAAGGACAGATGATCGTGTTCATTTGATGATGACATTTTGAGATTAATGAGATTAAC 3269  
 QY 1056 opProProProthCysProProthGluValGlnGluLeuMetGlnLeuGlyStrAlaProG 1076  
 Db 3270 AAGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3329  
 QY 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093  
 Db 3330 TGTAAATCAACGCCCTTCCTTACGATCTAGCTTCCTGAGTGCATCAATA 3381

## RESULT 4

AAAX80971 ID AAX80971 standard; cDNA; 4482 BP.

AC AAX80971;

DT 03-SEP-1999 (first entry)

XX Human JAK2 kinase encoding cDNA.

XX JAK2 kinase; arteriosclerosis; asthma; bronchitis; emphysema; psoriasis;  
 XX inflammatory bowel disease; inflammation; osteoarthritis; oncogenesis;  
 XX rheumatoid arthritis; septic shock; systemic lupus erythematosus;  
 XX leukaemia; human; ss.

OS Homo sapiens.

XX US5914393-A.

XX 22-JUN-1999.

XX 05-DEC-1995; 95US-0567508.

XX 05-DEC-1995; 95US-0567508.

XX (INCY-) INCYTE PHARM INC.

PI Coleman R, Stuart SG;

XX WPI; 1999-384188/32.

DR P-PSDB; AAI21698.

PT Polypeptides and amino acids useful for modulating human jak2 kinase  
 activity

PS Claim 2; Columns 23-28; 37pp; English.

XX This cDNA encodes a human JAK2 kinase polypeptide. Host cells transformed  
 CC with recombinant jak2 kinase nucleic acid are used for the recombinant  
 CC production of the protein. Purified JAK2 may be used to produce  
 CC antibodies or identify antagonists or inhibitors of JAK2. JAK2, anti-JAK2  
 CC antibodies and JAK2 antagonists or inhibitors may be used to treat.  
 CC prevent or diagnose conditions associated with altered or uncontrolled  
 CC Jak2 expression, e.g. arteriosclerosis, asthma, bronchitis, emphysema,  
 CC inflammatory bowel disease, inflammation, leukemia, oncogenesis,  
 CC osteoarthritis, psoriasis, rheumatoid arthritis, septic shock and  
 CC systemic lupus erythematosus.

XX Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other:

## Alignment Scores:

Pred. No.: 2,37e-190 Length: 4482  
 Score: 2621.00 Matches: 532  
 Percent Similarity: 66.03% Conservative: 193  
 Best Local Similarity: 48.45% Mismatches: 342  
 Query Match: 44.73% Indels: 31  
 DB: 20 Gaps: 14

US-09-397-967-16 (1-1099) x AAX80971 (1-4482)

QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProG 39  
 Db 447 TCAGATCTTCTAGGT-----GTATCTTTACCATCTCCCTGGGAATGTGAGGAGA 497  
 QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuGlyValAlaAlaAla 59  
 Db 498 TTATCTGACCTTCCATCTGGGAGATGTTGGAGAAATCTGTATTCCTCTCTAA 557  
 QY 59 sAlaCysGlyLeuLeuProValTyrHisSerLeuPheAlaLeuAlaThrGlyAspPhe 79  
 Db 558 AGCTTGtGGATCAACCTGTGTATCATATATATGTTGCTTAATGATGATCAAGAAAG 617  
 QY 79 rCysTrpPheProPheSerHisIlePheCysIleGluAspValAspThrGlnValLeu 99  
 Db 618 GATCTGTATCACCACCAACCTGTCTTCATATGATGATGATGATGATGATGATGATGAT 677  
 QY 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117  
 Db 678 CTACAGATTAAGATTTTACTTCTCTGTTGATGATGATGATGATGATGATGATGATGAT 737  
 QY 117 gPheGlyLeuArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisIle 137  
 Db 738 GCATGGAATATCTGAGGTGCTGAAGCTCCTCTTCTATGATGATGATGATGATGATGAT 797  
 QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMet 157  
 Db 798 CTTTGCTCAGCGCGCATGATTTTGTCATGATGATGATGATGATGATGATGATGATGAT 857  
 QY 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlu 177  
 Db 858 AACACAGAAAGATGCTTGGATGACAGTGTATGATATGATGATGATGATGATGATGAT 917  
 QY 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProPhe 197  
 Db 918 CGATCAAAACCCACTGGGCATCTATACCTATACCTATACCTATACCTATACCTATACCT 977  
 QY 197 rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgGlyIleArgGlyThr 217  
 Db 978 TATTCAGCAAAAGATCCAGTATCTATCTATCTATCTATCTATCTATCTATCTATCT 1037  
 QY 217 lValLeuAlaLeuLeuProGlyArgLeuProGlyArgProTyrAlaLeuMetAla 237  
 Db 1038 TCGCAGATTTTTCAGCATTCAGCAATGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1097  
 QY 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal 257  
 Db 1098 GTATCTTATAATCTGCAAACTGCAAGTCTGCTCTACACAGAAATTTGAAGTAA 1157  
 QY 257 yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs 273  
 Db 1158 AGAACCTGGAAGTGCTCCTCAGGTGAGAGATTTTTCGAACCATTTATTAATGAGAA 1217  
 QY 273 pAsnGlyIleProTyrSerSer-----AsnAspGluLe 284  
 Db 1218 CGGAGGAATTCAGTGCTCAAGGAAACATAAGAAAGTGAACACTGACAGAAAGCA 1277  
 QY 284 uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304  
 Db 1278 TTACAGTATATGCGATTTTCTTAATATATGATCTCAGTTTAAAGCAAGCAACCA 1337  
 QY 304 gValGlyProAlaGlyGlnHisArgLeuValThrValThrArgMetAspGlyHisIle 324  
 Db 1338 AGAGGGT--TCAATGAAAGCGAGTTGATCATATCATATGAAAGTGGTAAATAATCT 1394  
 QY 324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyr 344  
 Db 1395 GGAATTCATCTTACGTCATTAAGGAAAGCTTTGCTCTTCTGTCATTAATGATGATA 1454  
 QY 344 rPheArgLeuLeuCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLe 364  
 Db 1455 TTATAGATTAATCTGACATGACATCATATACCTGTAAAGATAGACCTCCAGCGCT 1514  
 QY 364 uLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLys 384

```

Db 1515 GCTTGAATAATATACAAAGCAACGTCATGGCCCAATTTCGATGGATTTGCCATTAGTAA 1574
      :|||||
Qy 384 sLeuLysAlaIaIaGlySerLeuProGlyThrTyrlLeuAlaArgSerProLysPly 404
      |||||
Db 1575 ACTGAACAAAGCGATTAATACAGCTGACCTGTATGACTTTCGATGACGCTCTTAAGACCTT 1634
      :|||
Qy 404 rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrlCysGly 424
      :|||
Db 1635 TAATAATATATTTTGTGACTTTTGTCTGTCGAGCCAGAAATGTCATGATTAATACACCTG 1694
      :|||
Qy 424 sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr 444
      :|||
Db 1695 TTTGATTTACAAATAAGAGAAATGAGATGACAACTCAGTGGGACAAAGAACTTCAG 1754
      :|||
Qy 444 gSerLeuArgGlyLeuLeuAlaIaIaCysTrpAsnSerGlyLeuAlaArgValAlaI 464
      :|||
Db 1755 CACTCTTAAAGATCTTTTGATTTGATTACCAAGATGAAACTGTTGCTCAGACAAATTAAT 1814
      :|||
Qy 464 vLeuTyrlLeuThrSerCysAlaIaProArgProLysGlyLysSerAsnLeuIleVala 484
      :|||
Db 1815 TTTCCAGATTACTAATAGCTGTCGCCCAAGCAAAAGTAATCAACCTTCTAGTCTT 1874
      :|||
Qy 484 IArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAlaIe 503
      :|||
Db 1875 CAGAAGCAATGGTGTTCGATGCTACCAACCTCACCAACATTACAGAGGCTACTCATAT 1934
      :|||
Qy 503 uThrGlnLeuSerPheIleThrIleProThrAspSerLeuGlyLysGlnAlaIleuL 523
      :|||
Db 1935 GAACCAAAATGCTGTTTCACAAATCAGAAATGAAAGATTGATTAATGAAGACCTTG 1994
      :|||
Qy 523 yHisGlySerPheThrLysIlePheArgGlySerArgArgGlyValAlaIaIaP---GlyG 542
      :|||
Db 1995 CCAGAGCACTTTTACAAAGATTTTAAAGCGCTACAGAGAGATGAGAGACTAGCTCA 2054
      :|||
Qy 542 uThrHisAspSerGlyValLeuLeuLysValMetAspSerArgHisArgAsCysMetG 562
      :|||
Db 2055 ACTGCATGAACAGAGATCTTTTAAAGCTCTGCTAAAGACACAGCACTATTCATCA 2114
      :|||
Qy 562 uSerPheLeuGlnAlaIaIaSerLeuMetSerGlnValSerTyrlProHisLeuValLeu 582
      :|||
Db 2115 GTCCTTCTTTGAAGCAGCAAGATGATGAGCAAGCTTCTCAGAACCTTGTGTTTAA 2174
      :|||
Qy 582 uHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnLysPheValTyrlLeuG 601
      :|||
Db 2175 TTATGGAGATGTGCTGCTGAGACAGAGATTTCTGCTCAGAGATTGTGTAATTTGG 2234
      :|||
Qy 601 yAlaIleAspMetTyrlLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeuG 621
      :|||
Db 2235 ATCACTAGATACATATCTGAAATAAATAAATTTGATTAATATATATGGAACCTTGA 2294
      :|||
Qy 621 nValThrLysGlnLeuAlaTyrlAlaLeuAsnTyrlLeuGlnLysGlyLeuProHisG 641
      :|||
Db 2295 AGTTCTTAACAGTGGATGGCCATGATCTTCTAGAACAAACACCCTTATTCATGG 2354
      :|||
Qy 641 yAsnValSerAlaArgLysValLeuLeuAlaArgGlnLysGly-----AspGlyAsnPr 659
      :|||
Db 2355 GAATTAATAGTCCAAATAATTTCTGCTATATCAGAGAAAGACAGAAAGCAAAATCC 2414
      :|||
Qy 659 oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGlnMetLe 679
      :|||
Db 2415 TCCCTTTCATCAAACTTACTGATCCGCGCATAGTATTTCACTTTTGGCAAAAGACATTC 2474
      :|||
Qy 679 uThrAspArgTyrIleProTrpValAlaProGlyCysLeuGlnLysGlnThrLeuCysLe 699
      :|||
Db 2475 TCAGAGAGGAATACCATGGGACACCACTGAATGCATGGAATCTTAATAATTTAAATTT 2534
      :|||
Qy 699 uGlnAlaAspLysTrpGlyPheGlyAlaIaIaThrTrpGlnLysAlaIaIaGlnArgGlyProAl 719
      :|||
Db 2535 GGCACACAGACAAATGGAGCTTTGGTACCACTTTGGGAAATTCGACGTGAGAGATTA 2594
      :|||
Qy 719 aHisIleThrSerLeuGlnProAlaLysLysLeuLysPheTyrlLysArgGlnLysGlnLe 739
      :|||

```

```

Db 2595 ACCTTAAGTCTCTGATTTCCAAAGAAAGCTACAAATTTATGAGATAGGCATCAGCT 2654
      :|||
Qy 739 uProAlaLeuLysTrpThrGlyLeuValaIaIaGlyLeuIleThrGlnCysMetAlaTyrlAspPr 759
      :|||
Db 2655 TCCCTGACCAAAAGTGGGACAGAAATTTACCAAACTTAAATAATTTGATGATTAATGAAACC 2714
      :|||
Qy 759 oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779
      :|||
Db 2715 AGATTTTCAGCCCTTCTTTCAGAGCCCATCATAGAGATCTTAACAGTTTGTTCATGCCAGA 2774
      :|||
Qy 779 rTyrlGlyLeuLeuSerAspProThrProGlyIleProSerProArg---AspGlyLeuCy 798
      :|||
Db 2775 TTATGACTATTAACAGAA---AATGACATGTTTACCAATTAAGAGATAGTGGCTTGGG 2831
      :|||
Qy 798 sValAlaGlyAlaGlnLeuTyrlAlaCysGlnAspProAlaIlePheGlnLysAlaArgHisLe 818
      :|||
Db 2832 GTTTTCTGCTGCC-----TTTGAAGACCGGGATGCTCACAGCTTGAAGAGACATTT 2885
      :|||
Qy 818 uLysTyrlIleSerLeuLeuGlyLysGlyAsnPheLysSerValGlyLeuCysArgTyrlAs 838
      :|||
Db 2886 GAATTTTCTACAGCAACTTGGGCAAGGTAATTTTGGAGTGTGAGATGTGCCCGGTATGA 2945
      :|||
Qy 838 pProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPr 858
      :|||
Db 2946 CCTCTACAGAGACACACTGGGAGAGTGTGCTCTTAAAGCTTCCAGCTAGATGTA 3005
      :|||
Qy 858 oAspGlnGlnArgAspPheGlnArgGlyLysGlnIleLeuLysAlaIleHisSerAspPr 878
      :|||
Db 3006 AGAGACCTTAAGAGACTTGAAGAGGAATGAATCTGAAATCTGAAATCTTACAGCATGACA 3065
      :|||
Qy 878 eIleValLysTyrlArgGlyValSerTyrlGlyProGlyArgGlnSerLeuAlaGlyLeuI 898
      :|||
Db 3066 CATTTTAAGATACAGAGAGAGTGTCTACAGTGTGCTGGGCTAAATTAATTAATAT 3125
      :|||
Qy 898 tGlnTyrlLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHis 917
      :|||
Db 3126 GGAATATTTTACCATATGGAAGTTTACAGAGATCTTCAAAACCTTAAGAACGATGGA 3185
      :|||
Qy 917 sThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlnMetGlnTyrlLeuGlyAl 937
      :|||
Db 3186 TCACATAAAGCTTCTGACGTACACATCTCAGATATGCAAGGATGAGAGTCTTTGTTGAC 3245
      :|||
Qy 937 aArgArgCysValHisArgAspLeuAlaIaIaArgAsnIleLeuValGlnLysArgGlnAlaI 957
      :|||
Db 3246 AAAAAGATATTCACAGAGATCTGCGACAGCAAGAAATATTTGTTGCGAGAGCAAGAACAG 3305
      :|||
Qy 957 sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlnLysAspTyrlTyrlVal 977
      :|||
Db 3306 AGTTAAATTTGGRGATTTTGGGTTAACCAAGTCTTGGCACACAGAACAAAGAAATCTATTA 3365
      :|||
Qy 977 lValArgGlnProGlyGlnSerProIlePheTrpTyrlAlaProGlyLysLeuSerAspAs 997
      :|||
Db 3366 AGTAAGAACACCTGCTGAAGTCCCATATCTTGTGATGCTCCAGAAATCAGTGAAGAGAG 3425
      :|||
Qy 997 nIlePheSerArgLysnSerAspValTyrlPheSerPheGlyValValLeuTyrlGlnLeuPheThr 1017
      :|||
Db 3426 CAAGTTTCTGTCGCTCAGATGTTTGGAGCTTGGAGTGTGCTGTATGATCACTTTTAC 3485
      :|||
Qy 1017 rTyrlCysAspLysSerCysSerProSerAlaGlnLysPheLysArgMetLeuGlyProGlnAr 1037
      :|||
Db 3486 ATTACATTGAGAGAGTAAAGTCCACAGCGGAAATTTATGCTGTATGATTTGGCAATGACA 3545
      :|||
Qy 1037 gGlnGlyProProLeuCys---ArgLeuLeuGlnLeuLeuAlaGlnLysArgTrpArgLeuPr 1056
      :|||
Db 3546 ACAAGACAGATGATCTGTTCCATTTTGATTAAGAACTTTTAAGAAATTAAGAAATTAAG 3605
      :|||
Qy 1056 oProProProThrCysProThrLeuValGlnLysLeuMetGlnLeuCysTrpArgProGly 1076
      :|||
Db 3606 AAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3665
      :|||
Qy 1076 uProHisAspArgProAlaPheAlaIleThrLeuSerProGlnLysAspProLeu 1093
      :|||
Db 3666 TGTAAATCAAGCCCTCTTGTAGGATCTTGTAGGATCTTGTAGGATCTTGTAGGATCTTGTAG 3717
      :|||

```

## RESULT 5

AAZ58947 ID AAZ58947 standard; cDNA; 4482 BP.

AAZ58947;

03-MAY-2000 (first entry)

Human JAK2 kinase (HUK2) encoding cDNA.

Janus family nonreceptor protein-tyrosine kinase-2; JAK2; HUK2; human; signal transduction; arteriosclerosis; asthma; bronchitis; emphysema; inflammatory bowel disease; leukemia; oncogenesis; osteoarthritis; psoriasis; rheumatoid arthritis; systemic lupus erythematosus; cytostatic; osteopathic; dermatological; antibacterial; septic shock; immunosuppressive; ss.

OS Homo sapiens.

PN US6019966-A.

PD 01-FEB-2000.

PE 19-NOV-1998; 9805-0196480.

PR 05-DEC-1995; 9505-0567508.

PA (INCY-) INCYTE PHARM INC.

PI Stuart SG, Coleman R;

DR WPI: 2000-146859/13.

DR P-PSDB; AAT77552.

Human Janus family nonreceptor protein-tyrosine kinases useful as diagnostic reagents and for preventing, diagnosing and treating diseases such as arteriosclerosis, asthma and leukemia.

Example 5; Fig 1A-F; 33pp; English.

This cDNA encodes a human Janus family nonreceptor protein-tyrosine kinase-2 (JAK2) polypeptide (HUK2). The JAK2 polypeptides may be used as diagnostic reagents as they react with a range of target proteins including growth hormone, prolactin, erythropoietin and cytokine receptors. They may also be used for the production of antibodies specific for JAK2, which may be used to inhibit its activity and prevent or treat disorders associated with over expression of JAK2. Conversely, the JAK2 polypeptide may be administered to supplement the patients own production and counter mutations that may lead to the expression of an inactive enzyme. The protein may also be used to screen candidate reagents for modulators of JAK2 function. The antagonists and antibodies bind to the JAK2 protein and prevent the transfer of high energy phosphate molecules, therefore blocking signal transduction. Disorders that may be treated by administration of JAK2 polypeptides, anti-JAK2 antibodies and the agonists and/or antagonists. Include arteriosclerosis, asthma, bronchitis, emphysema, inflammatory bowel disease, leukemia, oncogenesis, osteoarthritis, psoriasis, rheumatoid arthritis, septic shock and systemic lupus erythematosus.

Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other:

## Alignment Scores:

Pred. No.: 2,37e-190 Length: 4482  
 Score: 2621.00 Matches: 532  
 Percent Similarity: 66.03% Conservative: 193  
 Best Local Similarity: 48.45% Mismatches: 342  
 Query Match: 44.73% Indels: 31  
 DB: 21 Gaps: 14

US-09-397-967-16 (1-1099) x AAZ58947 (1-4482)

QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProG1 39

```

Db 447 TCCAGTCTTCAGT-----GTATCTTACCATTCCTGGGAATCTGAGCAGA 497
QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuGlyValArgAlaIaIy 59
Db 498 TTATCTGACCTTCCATCTCGGGAGATGTGTGAGAGAAATCTGTATCTGCTTCTTAA 557
QY 59 sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPhe 79
Db 558 AGCTTGTGATACACCTGTGTATCATATATGTCTTTAATGAGGAACAGAAAG 617
QY 79 rcYstPheProPheSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99
Db 618 GATCTGTATCCACCAACCATGTCTTCATATAGATGAGTACACAGGCATATGACT 677
QY 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117
Db 678 CTACAGAAATAGATTTCCTCTGCTTGTTGATTCAGTGGCAGCAACAGACCCATG 737
QY 117 gPheGlyLeuArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
Db 738 GCATGGAAATATCTCGAGGTCTGAGCTCTCTTGTATGACTTGTATCTTACT 797
QY 137 urPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157
Db 798 CTTTGCTCAGTGGCGCATATTGTGATGATGATGATGATGATGATGATGATGATGAT 857
QY 157 sGluGlnGlyIlePheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluG1 177
Db 858 AACACAGAGAAAGATGTCTTGGATGACAGTGTAGATATGATGATGATGATGATGAT 917
QY 177 nAlaGlnArgProGlyIleLeuLeuLysThrValSerTyrTyrLysAlaCysLeuProPhe 197
Db 918 CGATCAAAACCCACTGGCCATCTATACCTATACCTATACCTATACCTATACCTATAC 977
QY 197 rLeuArgAspValIleGlnGlyGlnAspPheValThrArgArgArgIleArgArgThrya 217
Db 978 TATTCGACAAAGATTCACAAACATCATATTTTGACAGAGAACGATAGATGACAGAT 1037
QY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaIy 237
Db 1038 TCCGAGATTTATTCAGCAATTCAGCCAAATGCAAAAGCCAGCCAGAACTTGAACCTAA 1097
QY 237 sTyrIleLeuAspLeuLysArgLeuHisProAlaIaIaThrGluThrPheArgValG1 257
Db 1098 GTATCTTATATATCTGAAACTCTGCAGTCTGCTTACACAGAAATTTGAAAGTAA 1157
QY 257 yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs 273
Db 1158 AGAACCTGGAAGTGTCTTCAGGTAGAGATTTTTCAAACCATTTATTAATCTGAAA 1217
QY 273 pAsnGlyIleProTrpSerSer-----AsnAspGluLe 284
Db 1218 CGGTGGAATTCAGTGCAGAGAGGAAACATTAAGAAAGTGAGACACTGACAGAACGAGA 1277
QY 284 urPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304
Db 1278 TTTCAGATTAATTCGATTTCTTAATATATATGATGATGATGATGATGATGATGAT 1337
QY 304 gValGlyProAlaGlyLysHisArgLeuValThrValThrArgMetAspGlyHisIleLe 324
Db 1338 AGAGGCT---TCAAATGAAGCCGAGTTGTATCATTCATTAAGCAAGATGTAAAAATCT 1394
QY 324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy 344
Db 1395 GGAATTCGAATTCATTCATTAAGGGAAGCTTGTCTTCTGTCATTAATTAATGATGATA 1454
QY 344 rPheArgLeuLeuLysAspSerArgHisTyrPheCysLysGlyValAlaProPheArgLe 364
Db 1455 TTATGATTAATTCAGATGACATCATCATTTACCTGTGTAACAAATAGACACCTCAAGCCGT 1514
QY 364 uLeuGlnGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLy 384

```

Db 1515 GCTTGAAATATACAAACCACTGTCATGCCAATTCATGATTTTGCATTAAGTAA 1574  
OY 384 sLeuYsAlaIaIaGlySerLeuProGlyThrTyrIleLeuArgSerProGlnAspTY 404  
Db 1575 ACTGAAACAGACGTAATCAGACTGTATGATGATTCGATGACGCTCTAAGACTT 1634  
OY 404 rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCY 424  
Db 1635 TAATAAATATATTTTGAATTTTGGCTGCAGCCGAAATAATTCATTGATTAACACTG 1694  
OY 424 sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAR 444  
Db 1695 TTTTATTAACAAATAATGAAATGAAAGATTCACACCTCAGTGGCAAAAGAACCTTCAG 1754  
OY 444 gSerLeuArgGlnLeuLeuAlaIaCysValGlnThrProLeuGlyLeuArgValAspGlyAlaI 464  
Db 1755 CAGCTTAAAGATCTTTTGAATTTTACACAGATGAAACCTGCTGCAGCAATATATAT 1814  
OY 464 aLeuTyrLeuLeuSerCysCysAlaProArgProLysGlyLysSerAsnLeuIleVala 484  
Db 1815 TTTCCAGTTTACTTAATGCTGTCCCAAGCCAAAGATTAATCAACCTTCTAGCTT 1874  
OY 484 LArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAlaIe 503  
Db 1875 CAGAACCAATGGTCTTCTGATGTACCAACCTCACCAACATTCACAGAGCCCTACTCAT 1934  
OY 503 uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGlyuThrHisGlnAsnLeuG 523  
Db 1935 GAACCAAAATGGTGTTCACAAATCAGAAATGAAAGATTTGATTAATTAAGAACCTTGG 1994  
OY 523 yHisGlySerPheThrLysIlePheArgGlySerArgGlyValValAsp---GlyG 542  
Db 1995 CCAAGGCACCTTTTAAAGATTTTAAAGCCTACAGAAAGAAAGTAAAGACCTACAGCTCA 2054  
OY 542 uThrHisAspSerGlnValLeuLeuLysValIleMetAspSerArgHisArgAsnGlyMetG 562  
Db 2055 ACAGCATGAACAGACAGACTTTTAAAGCTTGGATTAAGCACACGAGAACTTTCAGA 2114  
OY 562 uSerPheLeuGlnAlaIaIaSerLeuMetSerGlnValSerTyrProHisLeuValLeu 582  
Db 2115 GTCTTTCTTTGAAAGCAGCAGATGATGACAGACCTTCTCACAGACTTGGTTTTAA 2174  
OY 582 uHisGlyValCysMetAlaIaIaAsp---SerIleMetValGlnGluPheValTyrLeuG 601  
Db 2175 TTTATGGATGTGTCTGTCGAGACAGAAATATTCGTTACAGAGCTTGAATAATTTGG 2234  
OY 601 yAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrLeuG 621  
Db 2235 ATCAGTATGATACATATCTGAAAGAAATAAATTTGATTAATATTTATGCAAACTTGA 2294  
OY 621 nValIThrLysGlnLeuAlaIaIaLysAsnTyrLeuGlnAspLysGlyLeuProHisG 641  
Db 2295 AGTTGGTAAACACTGGCAGGCCCATGCTTTTCTAAGAAACACCTTATTCCTG 2354  
OY 641 yAsnValSerAlaArgLysValLeuLeuAlaArgGlyGlyGly-----AspGlyAsnPr 659  
Db 2355 GAATGATATGCCCCAAATATTTCTGCTTATCAGAGAAAGAACAGAGAACAGAAATCC 2414  
OY 659 oProPheIleLysLeuSerAspProGlyLysSerProThrValLeuSerLeuGlnMet 679  
Db 2415 TCCTTTCACTCAAACTAGTATGCTGCAATAGTATTACAGTTTCCAAAGACACTTCT 2474  
OY 679 uThrAspArgIleProThrValaIaIaProGlyCysLeuGlnGlnAlaGlnThrLeuCys 699  
Db 2475 TCGAGAGAGATATCAGATGGTACCACTGATGCAATTAATAATCTTAATATTT 2534  
OY 699 uGlnAlaAspLysTyrPheGlyAlaIaIaThrTyrPheGlnValPheGlnArgGlyProAl 719  
Db 2535 GGCACACCAATATGAGATTTTGTACCACTTGTGTGCAAAATCTGCAGTGGAGAGACATA 2594  
OY 719 aHisIleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGlnAspGlnGlyGln 739  
Db 2595 ACCTTAAGTGTGTGATTTCAAGAAAGTACAAATTTATGAGATGAGCATCAGCT 2654  
OY 739 uProAlaLeuLysTyrThrGlnLeuAlaIaIaLeuIleThrGlnCysMetAlaIaIaAsp 759  
Db 2655 TCCTGACCAAAAGTGGCAGACAAATAGCAAACTTATTAATTAATTTGATTAATGAC 2714  
OY 759 oGlyArgArgProSerPheArgAlaIleLeuAlaArgAspLeuAsnGlyLeuIleThrSer 779  
Db 2715 AGATTTCAGCCCTTCTTTCAGACCCATCATACAGATCTTAACAGTTTGTACTCCAGA 2774  
OY 779 pTyrGlnLeuLeuSerAspProThrProGlyIleProSerProArg---AspIleuLeuCY 798  
Db 2775 TTTATGAACCTTTTAACAGAA---AATGACATGTTACCAATATAGATATAGTCCCTGG 2831  
OY 798 sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGlnIuArgHis 818  
Db 2832 GTTTTCTGTGGCC-----TTTGAAAGACCGGATCTTCACACACTTTGAAGAGACATTT 2885  
OY 818 uLysTyrIleSerLeuLeuGlnGlyLysAsnPheGlySerValGlnLeuCysArgTyrAs 838  
Db 2886 GAATTTTCTACACCAACTTGGCAAGGTAAATTTTGGAGTGGAGATGTGCCGTATGA 2945  
OY 838 pProLeuGlyAspAsnThrGlyProLeuValaIaValLysGlnLeuGlnHisSerValPr 858  
Db 2946 CCTCTACAGAGACACACTGGGAGGTGGTCCGTAAAAAACCTTCACACATATGACTGA 3005  
OY 858 oAspGlnGlnArgAspPheGlnArgGlnIleGlnIleLeuLysAlaLeuHisSerAspH 878  
Db 3006 AGAGCACCTTAAGACCTTTGAAAGGAAATTTGAAATCCGAAATCCCTACAGCATGACA 3065  
OY 878 eIleValLysTyrFArgGlyValaSerTyrGlyProGlyFArgGlnSerLeuArgLeuIle 898  
Db 3066 CATTGTAAGGTACAAAGAGAGTGTGTACAGTGTGTGTCGGCTGAATCTTAATTAATTA 3125  
OY 898 tGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuH 917  
Db 3126 GGAATATTTACCATATGAGAGATTTACAGACATATCTTCAAAAACATTAAGAACGATGA 3185  
OY 917 sThrAspArgLeuLeuLeuPheAlaIaIaThrProGlnIleCysLysGlyMetGluTyrLeuG 937  
Db 3186 TCACATAAACCTTCTCAGTACATCTCAGATATGCAAGGTATGAGATATCTTGGTATC 3245  
OY 937 aArgArgCysValHisArgAspLeuAlaIaIaArgAsnIleLeuValGlnSerGlnAlaI 957  
Db 3246 AAAAAGGTATATCCACAGAGATCTGGCAACGAAATATATTTGTGTGGAGAACAGACAG 3305  
OY 957 sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVa 977  
Db 3306 AGTTAAATTTGARGATTTTGGGTAAACCAAGCTTGGCCACAAAGACAAAGATACTATA 3365  
OY 977 lValArgGlnProGlyGlnSerProIlePheThrPyrAlaIaIaProGlnSerLeuSerAsp 997  
Db 3366 AGTAAAGAACCTGTGGAAGTCCCATATTTCTGTATGCTTCAGAAATCTCAGAAATCAGAG 3425  
OY 997 nIlePheSerArgGlnSerAspValTyrPheGlyValValLeuTyrGlnLeuPheTh 1017  
Db 3426 CAAGTTTCTGTGGCTCAGATGTTTGGAGCTTGGAGGTGTCTGTATGAACTTTTTCAC 3485  
OY 1017 rTyrCysAspLysSerCysSerProSerAlaIaIaPheLeuArgMetMetGlyProGlnAR 1037  
Db 3486 ATACATTTGAAAGAGTAAAGTCCACACCGGAATTTATGCTATGATTTGGCAATGACA 3545  
OY 1037 gGlnGlyProProLeuCys---ArgLeuLeuGlnLeuLeuAlaGlnGlyArgArgLeuPr 1056  
Db 3546 ACAAGAGCAGATGATGCTTCATTTGATTAAGTAACTTTGAAAGATTAATGGAAGATTAC 3605  
OY 1056 oProProProThrCysProThrGlnValGlnIleLeuMetGlnLeuCysTTPAlaProG 1076  
Db 3606 ARGACCAAGTGGATCCCAAGATGATATATGATCATGACAGATCTGTGCAACATTA 3665  
OY 1076 uProHisAspArgProAlaPheAlaIaIaThrLeuSerProGlnLeuAspProLeu 1093  
Db 3666 TGTAAATACAGCCCTCTCTTATGAGATCTTGTGAGTGTGATCAATA 3717



```

RESULT 6
AAD24311
ID AAD24311 standard; DNA; 5117 BP.
AC AAD24311;
XX 07-MAR-2002 (first entry)
XX Human Jak2 (Janus kinase) DNA.
XX
XX Human: SOCS: suppressor of cytokine signalling; autoimmune disorder; Jak;
XX Janus kinase: signal transducer and activator of transcription; Stat;
XX rheumatoid arthritis; Wegener's granulomatosis; glomerulonephritis;
XX atopic dermatitis; multiple sclerosis; myasthenia gravis; vasculitis;
XX Crohn's disease; haemolytic anaemia; nephrotic syndrome; dermatological;
XX diabetes mellitus; thyroiditis; inflammatory bowel disease; nephrotropic;
XX immunosuppressive; antiinflammatory; antineumatic; antiarthritic;
XX hepatotropic; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 478..3876
XX /tag="a
XX /product="Human Jak2 protein"
XX
XX WO200179555-A2.
XX
XX 25-OCT-2001.
XX
XX 13-APR-2001; 2001WO-US12131.
XX
XX 14-APR-2000; 2000US-0549654.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Hancock WW, Ozkaynak E;
XX
XX WPI: 2002-034368/04.
XX P-PSDB: AAE15179.
XX
XX Monitoring transplant acceptance or autoimmune disease, useful e.g. for
XX assessing therapy, comprises measuring levels of Stat or their
XX inhibitors -
XX
XX Example: Fig 8; 218pp; English.
XX
XX The present invention relates to a method for monitoring acceptance of a
XX transplant or an autoimmune disease in a mammal. The method comprising
XX determining the amount of at least one of Stat4 (signal transducer and
XX activator of transcription), Stat6, SOCS1 (suppressor of cytokine
XX signalling) or SOCS3 mRNA or protein in a sample of the transplant, taken
XX from the host or an affected tissue sample. Stats are activated by
XX receptor-associated Janus kinases (Jaks) which include Jak1, Jak2, Tyk2,
XX Jak3. The method is used to determine whether acceptance of a transplant
XX has been induced or to determine if autoimmune disorders (systemic
XX lupus erythematosus, glomerulonephritis, rheumatoid arthritis, Wegener's
XX granulomatosis, chronic active hepatitis, atopic dermatitis, multiple
XX sclerosis, myasthenia gravis, haemolytic anaemia, nephrotic syndrome,
XX thyroiditis, diabetes mellitus, Crohn's disease, inflammatory bowel
XX disease and vasculitis) are being treated successfully and may be used
XX to adjust treatment regimes. The present sequence is human Jak2 DNA.
XX
XX Sequence 5117 BP; 1623 A; 927 C; 1103 G; 1464 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,84e-190 Length: 5117
Score: 2621.00 Matches: 533
Percent Similarity: 66.12% Conservative: 193
Best Local Similarity: 48.54% Mismatches: 341
Query Match: 44,738 Gaps: 31
DB: 24 Gaps: 14

```

```

US-09-397-967-16 (1-1099) x AAD24311 (1-5117)
QY 20 SerSergluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGly 39
Db 585 TCCAGTCTTCAGGT-----GTATCTTTACCATTCCTTGGAAGATTCGAGCAGCA 635
QY 39 natgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuGlyValArgAlaIaIa 59
Db 636 TTAATGACCTTTTCACATCTGGGAGATATTCGAGAAAGAAATCGATATTCCTCTTA 695
QY 59 salAcysGlyIleLeuProValTyrHisSerLeuPheAlaIleValAlaThrcIuAspPhe 79
Db 696 AGCTTGATGATACACACCTGGATCATATATCTTTGTTAATGAGTGAACAGAAAG 755
QY 79 rGystrPheProProSerHisIlePheCysIleGluAspValAspPheIleValIleVal 99
Db 756 GATCTGATTCACCAACCAACCATGCTTCATATAGATGAGCAACGCAATATGACT 815
QY 99 lTyArgLeuArgPheTyrPheProAspTyrPhe-----GlyLeuGluThrCysHisAr 117
Db 816 CTACAGATATAGATTTTACTTCTCGTGGTATTCGAGGAGCAACAGAGCCTATCG 875
QY 117 gPheGlyLeuArgIlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisIe 137
Db 876 GCATGGAATATCTCGAGAGTGTGAGAGCTCCTCTTGATGACTTTTGCATGCTTACT 935
QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetIy 157
Db 936 CTTCCTCAGTGGCGGCGCATGATTTTGTGCACGAGTGGATTAATACCTGCTGACTGA 995
QY 157 sGluGlnGlyIlePheLeuSerIleAlaValLeuAspLeuAlaGlnMetAlaArgGluG 177
Db 996 AACACAGGAAGAAATGCTTGGATGCGCAGTGTATGATATGAGTAAGCAAGAAAGAAA 1055
QY 177 nAlaGlnArgProGlyGlyLeuLeuLysThrValSerTyrIleLysAlaCysLeuProPse 197
Db 1056 CGATCAAAACCCACTGGCCATCTATACCTATACGCTACAGACATTTCTTACCAAAATG 1115
QY 197 rLeuArgAspValIleGlnGlyGlnAspPheValThrArgArgIleArgArgIleVal 217
Db 1116 TATTCGACCAAAAGATCCAAAGACTATCATATTTTGACAGAAAGCAATAGGTACAGATT 1175
QY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProIlyrAlaLeuMetAlaIy 237
Db 1176 TCGCAGATTATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAG 1235
QY 237 sTyrlIleLeuAspLeuGluArgLeuHisProAlaIaIaThrIleGluThrPheArgValG 257
Db 1236 GTATCTTTAATATGGAACCTGCTGACCTTGTACACAGAAATTTGAAGTAA 1295
QY 257 yLeuProGlyAla-----GlnGluGlnProGlyLeuLeuArgValAlaIleGlyAs 273
Db 1296 AACACCTGGAAGTGTCTCTTCAGTGCAGAGATTTTTCACACATATATTAATGCTGAAA 1355
QY 273 pAsnGlyIleProIlyPseSer-----AsnAspGluIle 284
Db 1356 CGTGGAAATTCAGTGTGTAAGAGGAAACATTAAGAAAGTACAGACATGAGAAACAGA 1415
QY 284 uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304
Db 1416 TTTACAGTTATATTCGATTTTTCATATATATATATATATATATATATATATATAT 1475
QY 304 gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspArgIleHisIle 324
Db 1476 AGAGGCT---TCAAATGAAGCCGAGTTGTAATTCATATCAATGAAGATGGTAAAAATCT 1532
QY 324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaIleValAspGlyIy 344
Db 1533 GGAATATGAACTTACGCTATTAAGGAAACCTTGCTTTCGTCATTAATATGATGAGATA 1592
QY 344 rPheArgLeuIleCysAspSerArgHisTyrPheCysIlyGluValAlaIleProArgIle 364
Db 1592 rPheArgLeuIleCysAspSerArgHisTyrPheCysIlyGluValAlaIleProArgIle 364

```



Db 1593 TTATACATTAACTGCAGATGCACATCATTAACCTGTGTAAAGAGTAGACCTCCAGCCGT 1652  
Qy uleuGLUGLUAlaAspValCySHISGLYProIleThrLeuAspPheAlaIleHisLy 384  
Db 1653 GCTTGAATAATATACAAAGCACTGTGATGCCCAATTTGCATGATGATTTGGCATTTGTA 1712  
Qy sleuLyAlaIaGLYSerLeuProGlyThrTyrIleLeuArgSerProGlnAsp 404  
Db 1713 ACTGAAGAAAGCAGATATACAGCTGATGTACTGTACCTGCATGCAGCTTAAGCACTT 1772  
Qy rAspSerPheLeuThAlaCysValGlnThrProLeuGlyProAspTyrIleGlyCy 424  
Db 1773 TAATTAATATTTTGTGATTTGCTGTGTCGAGCAAAATGTCATGATTAATTAACACTG 1832  
Qy sleuIleArgGlnAspSerSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr 444  
Db 1833 TTTCGATTACAAAATGACAGATAGACATACACCTCAGTGGGACAAAGAACACTTCAG 1892  
Qy gSerLeuArgGlnLeuLeuAlaIaCysThrPasnSerGlyLeuArgValAspIlyAlaAl 464  
Db 1893 CAGTCTTAAAGATCTTTGCAATGTTCACAGATGAGAACCTTCGCTCAGACAAATTAAT 1952  
Qy AleuTyrLeuThrSerCysCysAlaProArGProLySGlyLeuSerAsnLeuIleValVa 484  
Db 1953 TTTCAGATTACTAAATGCTGTCCCAAGGCCAAAGATAAATCAAACTCTTACTGCTT 2012  
Qy lArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAlaLe 503  
Db 2013 CAGAACGATGCTGTTTGCATGTACCAACCTCACACATTCACAGGCTCTCATAT 2072  
Qy uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGlyThrPheGlnAsnLeuG 523  
Db 2073 GAACCAAAAGCTGTTTCAAAATCAGAAATGCAAGATTTGATTAATGAAGACCTTGG 2132  
Qy yHISGLYSerPheThrLyIlePheArgLySerArgArgValAlaAsp---GlyLeu 542  
Db 2133 CCAGGACCTTTTACAAAGTTTTTAAAGCCGTACAGAGAGATGAGACAGCTGCTCA 2192  
Qy uThrHisAspSerGlnValLeuLeuLyValMeCAspSerArgHisArgAsnCysMetG 562  
Db 2193 ACTGACATGAACAGAGATCTTTTAAAGTTTGCATTAAGCAGACAGAACTATTCAGA 2252  
Qy uSerPheLeuGlnAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLe 582  
Db 2253 GCTTTCTTTGAGCAGCAGATATGATGAGCAAGCTTTCTCACACACTTTGGTTTAA 2312  
Qy uHISGLYAlaCysMetAlaGlyAsp---SerIleMetValGlnGlyPheValTyrLeuG 601  
Db 2313 TTATGAGATATGCTGTGAGACAGACAGATATTCGTTCAAGACTTTGTAAATTTGG 2372  
Qy yAlaIleAspMetTyrLeuArgLyValGlyHisLeuValSerAlaSerTyrPheLeuG 621  
Db 2373 ATCAGTATGATATCTGTAATAAAAGAAATTAATGTAATATATGTAAGAACTTGA 2432  
Qy nValThrLySGlnLeuAlaIaTyrAlaLeuAsnTyrIleGlnAspLyGlyLeuProHisG 641  
Db 2433 AGTTGCTTAAACAGTTGGCTGGCCATGCAATTTTCTAGAAGAAACACCCCTTAATGATG 2492  
Qy yAsnValSerAlaArgLyValLeuLeuAlaArgGlyGlyLy-----AspGlyAsnPr 659  
Db 2493 GAAATGATATGCCAAATAATTTCTGTTATCAGAGAAAGACAGAACAGAGAAATTC 2552  
Qy oProPheIleLyLeuSerAspProGlyValSerProThrValLeuSerLeuGlyMetLe 679  
Db 2553 TCCCTTCATCAAACTAGATCTGCGCATATGATTAACAGTTTCCCAAGCACTTCT 2612  
Qy uThrAspArgIleProThrPValAlaProGlyCysLeuGlnGlyAlaGlnThrLeuGly 699  
Db 2613 TCAGAGAGAAATACATGCTGACACCTGAATGCATGTAATCCATAAAATTTAAATTT 2672  
Qy uGlnAlaAspLySTrPGLYpHeGlyAlaIaThrThrTrpGlyValAlaPheGlnArgGlyProAl 719  
Db 2673 GGCAACAGACAAATGAGATTTTGGTACCACTTTTGGGAAATCTGCAGTGGAGAGATTA 2732

Qy 719 aHisIleThrSerLeuGluProAlaLySLeuLySLeuPheTyrGlnAspGlnGlyIle 739  
Db 2733 ACCTTAAGTGCCTGTGATTTCTCAAGAAAGATTAATTTTATGAGATGAGCATAGCT 2792  
Qy uProAlaLeuLySTrPThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAsp 759  
Db 2793 TCCGTGACCAAAAGTGGCGCAGAAATTTACCAAACTTAAATTAATTTGATGATTAACACC 2852  
Qy oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779  
Db 2853 AGATTTTCAGCCCTTCTTACAGACCATCATACAGATCTTAACAGTTTGTACTCCACA 2912  
Qy pTyrglyLeuLeuSerAspProThrProGlyIleProSerProArg---AspGlyLeu 798  
Db 2913 TTATGACTATTTTACAGAA---AATACATGTTTACCAAAATGAGATGAGTGGCCCTGGG 2969  
Qy sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGlyArgHisLe 818  
Db 2970 GTTTCTGGTGCC-----TTTGAAGACCGGATCCTACACAGTTTGAAGAGACATTT 3023  
Qy uLySTyrIleSerLeuLeuGlyLySGlyAsnPheGlySerValGlnLeuCysArgTyrAs 838  
Db 3024 GAAATTTCTACAGCAACTTGGCAAGGTAATTTTGGAGTGTGAGATGTGCGGCTATGA 3083  
Qy pProLeuGlyAspAsnThrGlyProLeuValAlaValLySGlnLeuGlnHisSerValPr 858  
Db 3084 CCCTCTACAGACAAACACTGGGAGGTGCTGCTGTAATAAAACCTTCACACTATGCTCA 3143  
Qy oAspGlnGlnArgAspPheGlnArgGlyIleGlnIleLeuLyAlaLeuHisSerAspPh 878  
Db 3144 AGAGACACTTACAGACATTTGAAAGGAAATTTGAAATTCCTGAAATCCCTCAGCATGACA 3203  
Qy eIleValLySTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgValMe 898  
Db 3204 CATGTAAAGTACAAAGGAGTGTGTACAGTGTGCTGTGCGGCTGAATCTGAATTAATTA 3263  
Qy tGlyTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuGln 917  
Db 3264 GGAATTTTACCATATGAGAAATTTACAGACATATCTTCAAAAACATAAAGAACGGATACA 3323  
Qy sThrAspArgLeuLeuPheAlaTrpGlnIleCysLySGlyMetGlyTyrLeuGlyAl 937  
Db 3324 TCACATAAACCTTGTCCACATCATCTGACATTCACAAAGGATATGAGATATCTTGTGAC 3383  
Qy aArgArgCysValHisArgAspLeuAlaIaArgAsnIleLeuValGlnSerGlnAlaH 957  
Db 3384 AAAAGGCTATTCACAGGAGATCTGCCAACGAAATATATGTTGGTGGACAGACAGACAG 3443  
Qy sValLySLeuAlaAspPheGlyLeuAlaLySLeuLeuProLeuGlyLyAspTyrTyrVa 977  
Db 3444 AGTTAAATTTGAGATTTTGGGTTAACCAAGCTTGGCCACAAAGCAAAAGATATGATA 3503  
Qy yValArgGlyProGlyGlnSerProIlePheTrpTyrAlaProGlySerLeuSerAspAs 997  
Db 3504 AGTAAAGAACCTGTGGAAAGTCCCATATTTCTGTATCTGCCAGAAATTCACAGACAG 3563  
Qy nIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheH 1017  
Db 3564 CAAGTTTCTGTGGCCTCAGATGTTTGGAGCTTTGGAGTGTCTGTATGAACCTTTTAC 3623  
Qy rTyrglyAspLySLeuSerCysSerProSerAlaGlnPheLeuArgMetMetGlyProGly 1037  
Db 3624 ATACATTTGAGAGAGTAAAGTCCACACAGGATTTATGCTATGATTTGGCAATGACAA 3683  
Qy gGlyGlyProProLeuCys---ArgLeuLeuGlnLeuLeuAlaGlnGlyArgArgLeuPr 1056  
Db 3684 ACAAGACAGATGATGCTGTTCCATTTGATTAAGACTTTTGAAGAAATATGAGAAATTAAC 3743  
Qy oProProProThrCysProThrGlnValGlnIleuMetGlnLeuGlySTrPAlaProG 1076  
Db 3744 AAGACCAAGATGATGCCACGATGAGATCATATGATCATGACAGAAATGCTGGAAACAATA 3803

OY 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093  
 DB 3804 TGTAAATCAACGCCCTTACGATCTAGCTTCTGAGTCAATA 3855

RESULT 7  
 AA085412  
 ID AA085412 standard; cDNA; 3629 BP.

XX  
 AC AA085412;  
 DT 05-OCT-1995 (first entry)  
 XX  
 DE Murine JAK2 kinase coding sequence and flanking regions.  
 XX  
 KW JAK family; protein tyrosine kinase; cytokine receptor; mouse;  
 KM phosphorylation; signal transduction; activation; ss.  
 OS Mus musculus.

XX  
 FH Key Location/Qualifiers  
 FT CDS 94..3483  
 FT  
 FT /\*tag= a  
 FT /\*product= Jak2\_kinase  
 FT 554..556  
 FT /\*tag= b  
 FT /\*note= "Published partial sequence of Jak2 cDNA  
 FT (Harpur et al., Oncogene 7:1347-1353(1992))  
 FT has CCC at this position"  
 FT 1089..3579  
 FT /\*tag= c  
 FT /\*note= "Published partial sequence of Jak2 cDNA  
 FT (Harpur et al., Oncogene 7:1347-1353(1992))  
 FT differs from AA085412 in having the  
 FT nucleotides shown in brackets at the following  
 FT positions: 1089(T), 1103(C), 1114(G), 1119(G),  
 FT 1122(C), 1128(C), 1131(G), 1134(G), 1137(C),  
 FT 1140(G), 1143(G), 1146(C), 1188(T), 1194(G),  
 FT 1230(A), 1245(T), 1260(T), 1266(C), 1272(T),  
 FT 1275(A), 1293(T), 1305(T), 1323(C), 1341(A),  
 FT 1344(A), 1359(G), 1365(A), 1368(T), 1374(T),  
 FT 1401(C), 1413(C), 1431(T), 1458(A), 1476(G),  
 FT 1488(T), 1511..1512(GT), 1578(C), 1590(T),  
 FT 1593(T), 1602(G), 1623(T), 1642(G), 1657(C),  
 FT 1728(G), 1743(C), 1755(C), 1770(A), 1809(G),  
 FT 1816(G), 1821(C), 1857(A), 1878(T), 1935(A),  
 FT 1938(A), 1963(T), 1974(G), 2025(T), 2055(G),  
 FT 2079(C), 2082(C), 2085(C), 2253(A), 2259(G),  
 FT 2283(A), 2285(C), 2433(G), 3453(C), 3579(C)."

FT conflict 2226  
 FT /\*tag= d  
 FT /\*note= "location of a 7 amino acid insert detected  
 FT in previous studies, but not in the present  
 FT study"  
 FT 3595..3619  
 FT /\*tag= e  
 FT /\*note= "three extra nucleotides (all A's) were  
 FT present in previous studies at positions  
 FT 3595, 3598 and 3619 in the 3'-UTR"

FT conflict  
 FT 2226  
 FT /\*tag= d  
 FT /\*note= "location of a 7 amino acid insert detected  
 FT in previous studies, but not in the present  
 FT study"  
 FT 3595..3619  
 FT /\*tag= e  
 FT /\*note= "three extra nucleotides (all A's) were  
 FT present in previous studies at positions  
 FT 3595, 3598 and 3619 in the 3'-UTR"

PN MO9503701-A.  
 XX  
 XX 09-FEB-1995.  
 PD  
 XX 29-JUL-1994; 94MO-US08676.  
 PF  
 XX 29-JUL-1993; 93US-0097997.  
 PR  
 XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 PA  
 XX Inle JN, Quelle FW, Silvenoinen O, Witthuhn BA;  
 PI WPL: 1995-081950/11.  
 XX P-PSDB: AAR70830.

XX  
 PT Inhibiting a cellular response to a cytokine by inhibiting Jak  
 PT kinase - to treat diseases caused by excessive response to  
 PT cytokine, e.g. erythrocytosis and other cellular proliferative  
 PT diseases  
 XX  
 PS Claim 29; Fig 1; 167pp; English.  
 XX  
 CC Inhibiting the activity of a jak kinase (pref. Jak1, Jak2, Jak3 or  
 CC Tyk2) in a eukaryotic cell is claimed as a method of inhibiting the  
 CC biological response of that cell to a cytokine (not IL-3 or  
 CC erythropoietin). The present sequence (murine JAK2 kinase) encodes a  
 CC 1129 amino acid protein which includes an epitopic sequence at  
 CC amino acid positions 758-776. Antibodies which selectively bind the  
 CC epitope are able to bind Jak2 without interfering with the activity  
 CC of the kinase. Such antibodies are claimed and are useful for  
 CC detecting and extracting Jak2. The 71 nucleotide differences noted  
 CC between the present sequence and the published partial sequence  
 CC result in 9 amino acid changes.

XX  
 SQ Sequence 3629 BP; 1142 A; 733 C; 842 G; 912 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,33e-190 Length: 3629  
 Score: 2619.50 Matches: 537  
 Percent Similarity: 65.50% Conservative: 190  
 Best Local Similarity: 48.38% Mismatches: 350  
 Query Match: 44.70% Indels: 33  
 DB: 16 Gaps: 15

US-09-397-967-16 (1-1099) x AA085412 (1-3629)  
 OY 4 ProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerGluAla 23  
 DB 145 CCTGTACATCAGATGATGATGATTCCTGGAAGTGCTAATTCGTGAGACAGATAGACCA 204  
 OY 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArgLeuSerPhe 43  
 DB 205 GTCCCTCAAGTGTATCTGTACCATTCCTTGGCAAGCGTAGGAGATATCAAGTTT 264  
 OY 44 SerPheGlyAspPyrLeuAlaGluAspLeuCysValArgAlaAlaValAlaCysGlyIle 63  
 DB 265 CCAAGTGGAGAGTATGTTGCAAGAAATTTGTGTGGCTGTTTAAAGCTTGCGATTT 324  
 OY 64 LeuProAlaThrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83  
 DB 325 ACGGCTGTATCAATATATGTTGCGTTATGAGTGAACCAAGAGATCTGCATACCA 384  
 OY 84 ProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg 103  
 DB 385 CCCAATCATGTGCTTCCACATATGACGATCAACAGGATGACATATCTACAGATTAAGG 444  
 OY 104 PheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAspPheGlyLeuArg 121  
 DB 445 TTCTACTTCCTCATGTGTACTAGTACGACGACGACAAACCTTACATAGAGGTGTC 504  
 OY 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141  
 DB 505 CGTGAGCGCTGAAGCTCCCTGCTGTGATGATCTTGTATGATGCTTACCTTTGTCACAGG 564  
 OY 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 161  
 DB 565 CGGCATGATTTTGTTCACGATGATTAAGTAAAGTACCTGTGACTCATGAACTCAGAGACG 624  
 OY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArgPro 181  
 DB 625 TGTCTGGGATGCGCGATTGTAGACATGATGAGANTAGTAGAAGAAAGACAGACTCCA 684  
 OY 182 GlyIleLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201  
 DB 685 CTGGCTGTCTAATCTGTGACGACACAAACATTTTACCAAGAGTGGCTGCAGCGAAG 744  
 OY 202 IleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAlaLeu 221



```
QY 919 AspArgLeuLeuLeuPheAlaTpgGlnIleCysLysGlyMetGluTyrLeuGluAlaArg 938
    :::::::::: ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 2944 AAAAACTTCTCAATACATCATCTCAAGATGAGAGGCGATGATATCTTGATACAAA 3003
QY 939 ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHisVal 958
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3004 AGGTATATCCACAGGGACCTGGGCACAGACATATTGGTGGAAATGAGAACAGGGTT 3063
QY 959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGluLysAspTyrTyrValVal 978
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3064 AATAATAGACACTTCGATTAACCAAGCTTGCCGACAGCAAAAGATACTCAAGATA 3123
QY 979 ArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAsnIle 998
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3124 AAGGAGCCAGGCGGAAACCCCATATCTGTACCCACCTGAATCTTGACGAGAGCAAG 3183
QY 999 PheSerArgGlnSerAspValTyrPheGlyValValLeuTyrGluLeuPheTyrTyr 1018
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3184 TTTTCTGTGGCTCAGATGTGTGTGAGCTTGAGCTCTTATACGAACTTTTCACATAC 3243
QY 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu 1038
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3244 ATCGAGAGAGTAAAGTCCACCCGTAATTTATGCGAATGATGGCAATGATGAACAA 3303
QY 1039 GlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuPro 1057
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3304 GGGCAAAATGATGTGTCTTCATTGTAGAGCTACTGAGAGCAACGAGAAAGATCCCAAG 3363
QY 1058 ProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTyrAlaProGluPro 1077
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3364 CCAGAGAGATGCCAGATGAGATTTATGTATGATCATGACAGAGTGTGAGAACAAATGTG 3423
QY 1078 HisAspArgProAlaPheAlaThrLeuSer 1087
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3424 AGCCAGCGTCCCTCTTCAGGGACCTTTCG 3453

RESULT 8
AAC66244
ID AAC66244 standard; cDNA; 3629 BP.
AC
AAC66244;
XX
DT 19-FEB-2001 (first entry)
XX
DE Jak2 polynucleotide sequence.
XX
KW Jak3; kinase; cytokine; cellular response; inhibition; jak2; ss;
KW cell proliferation; erythrocytosis.
XX
OS Unidentified.
XX
PN US6136595-A.
XX
PD 24-OCT-2000.
XX
PE 18-JUN-1996; 96US-0665574.
XX
PR 29-JUL-1994; 94US-0282012.
PR 29-JUL-1993; 93US-0097997.
PR 09-SEP-1993; 93US-0118968.
XX
PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI Silvenoinen O, Witthuhn BA, Ihle J;
XX
DR WPI; 2000-686080/67.
XX
P-PSDB; AAB35719.
XX
PT New DNA encoding Jak3 kinase is useful as cytokine regulator for
XX treating cell proliferation
XX
XX Example 1; Fig 1; 100pp; English.
```

```
XX
CC This invention relates to DNA encoding a murine Jak3 protein. The amino
CC acid sequence of the Jak3 protein is given in AAB35715. The jak family
CC of kinases are involved in the cellular response to the binding of
CC cytokines to their respective receptors. Jak3 kinase mediated activation
CC of some cytokines through their phosphorylation in response to
CC cytokine-receptor binding. Inhibiting the activity of Jak3 kinase (at the
CC nucleic acid level with antisense sequences or ribozymes, or at the
CC protein level with antibodies, kinase inhibitors etc.) is used to treat,
CC or diagnose, diseases caused by excessive secretion of certain cytokines,
CC e.g. excessive cell proliferation such as erythrocytosis. Alternatively,
CC the Jak3 protein and polynucleotide can be used to treat conditions
CC associated with defective Jak3 activity. The DNA sequence can be used to
CC produce recombinant Jak3 and this used to raise antibodies useful as
CC specific inhibitors or to detect or isolate Jak3 without interfering
CC with its enzymatic activity. The present sequence represents cDNA
CC encoding a Jak2 protein used in the isolation and characterisation of the
CC Jak3 protein of the invention.
XX
SQ Sequence 3629 BP; 1142 A; 732 C; 843 G; 912 T; 0 other:
```

```
Alignment Scores:
Pred. No.: 2,33e-190 Length: 3629
Score: 2619.50 Matches: 537
Percent Similarity: 65.50% Conservative: 190
Best Local Similarity: 48.38% Mismatches: 350
Query Match: 44.70% Indels: 33
DB: 21 Gaps: 15
```

US-09-397-967-16 (1-1099) x AAC66244 (1-3629)

```
QY 4 ProSerGlnGluThrProLeuIleProGlnArgSerCysSerLeuSerSerGluAla 23
    :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 145 CCTGTACATCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 204
QY 24 GlyValAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArgLeuSerPhe 43
    :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 205 GTCCCTCAAGTATGTCGTACCATCTCTTGGGCAACTGAGAGAGAGATTCGAACTTT 264
QY 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysGlyIle 63
    :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 265 CCAGTGGAGAGTATGTCGAGAGAAATTTGTGCTGCTTAAAGCTTGATGAT 324
QY 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTyrPhePro 83
    :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 325 ACGCCTGTGTATCAATATATGTCCTTATATGATGATGATGATGATGATGATGATGAT 384
QY 84 ProSerHisIlePheCysIleGluAspValAspThrGlnValTyrArgLeuArg 103
    :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 385 CCCAATCATGTGCTTCCACATAGACGAGTCAACGAGCATGACATACCTACAGGATAAG 444
QY 104 PheTyrPheProAspTyrPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg 121
    :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 445 TTCTACTTCCCTCATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 504
QY 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141
    :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 505 CGTGGGGCTGAGACCTCCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 564
QY 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu 161
    :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 565 CGGCATGATTTTGTTCACGAGTGAATAAGTACCTGATGATGATGATGATGATGATGATGAT 624
QY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArgPro 181
    :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 625 TGCTGTGGAGTGGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 684
QY 182 GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201
    :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
Db 685 CTGGCTGTCTAACTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 744
QY 202 ILeGlnGlyGlnAspPheValThrArgArgArgIleArgValThrValLeuAlaLeu 221
    :::::::::: :::::::::: :::::::::: :::::::::: ::::::::::
```

Db 745 ATCCAGACTATTCACATTTAACCCGGAAGCAATTCAGATTCGACATTCATT 804  
Qy 222 LeuProCysGlyArgLeuProGlyArgProThrValAlaLeuMetAlaIleuValIleuLeuAsp 241  
Db 805 CAGCAATTCAGTCAATGTAAGAACCCACACGCCAGAACCTTAACCTTAAGTATCTTAATAAC 864  
Qy 242 LeuGluArgLeuHisProAlaAlaIleThrGluThrPheArgVal 256  
Db 865 CTGGAAACCCCTGCAGTCCTCTTACACAGAACAGTTTGAAAGTAAGAAATCTGCAGAA 924  
Qy 257 GlyLeuProGlyAlaGluGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle 276  
Db 925 GGT---CCTTCAGCTGAGGAGATTTTGCACCACTTAATTAACCTGGAACCGGTGGAAT 981  
Qy 277 ProTrpSerSer 1041  
Db 982 CAGTGGTCAAGAGGGAACATTAAGGAAGTGAGCACTGCACACAGACGACGATCACTTA 1041  
Qy 288 PheCysAspPheProGluIleValAlaSerIleAsnGlnAlaProArgValGlyPro 307  
Db 1042 TATTGTATTTCCCTGATATTAATGATGTCGATTAAGCAACCAACGAG---GAATGC 1098  
Qy 308 AlaGlyGluHisArgLeuValIleThrValIleArgMetAspGlyHisIleLeuGluAlaGlu 327  
Db 1099 TCAGATGAAGTAGTAAGATGTAAGTCTGCATTAACAAAGATGAAGATTTTGGAGATAGAA 1158  
Qy 328 PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyIleThrPheArgLeu 347  
Db 1159 CTTAGCTCATTAATAAGAGCCTTGTCATTCGTCATTAATTAAGCGGATTAACAGACTA 1218  
Qy 348 IleCysAspSerArgHisArgIlePheCysGlyGluValAlaProProArgLeuGluGlu 367  
Db 1219 ACTGCGGATGCGCACCATTAACCTCTGCAAGAGGTGGTCCGCCACCTGCTGCGAGAAC 1278  
Qy 368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisIleuLeuValAla 387  
Db 1279 ATACACAGCAACTGCCAGCGCCCAATATCAATGATTTTGCATTAAGCAACTTAAGAAAG 1338  
Qy 388 AlaGlySerLeuProGlyIleThrIleLeuArgArgSerProGlnAspTrpAspPhe 407  
Db 1339 GCGGGATACCAAGCTGATATATGTCATGATGACGCGCTTAAGAGACTTCAACAAATATAC 1398  
Qy 408 LeuLeuThrAlaCysValAlaIleThrProLeuGlyProAspTrpIleGlyCysIleuLeuArg 427  
Db 1399 TTTTCGACCTTTGCTGTTGAGGAGAAATGTCATTAATTAACACTGTTTGAATACG 1458  
Qy 428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447  
Db 1459 AAGAAATGAAATGAGAAATACAACTCAGCGGACTTAAGAGAACTTCAGTAACCTTAAG 1518  
Qy 448 GluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuIleu 467  
Db 1519 GACCTTTTGAATTTGCTACCAAGATGGAACCTGCGCTCAGACAGATATCTTCCAGTTT 1578  
Qy 468 ThrSerCysCysAlaProArgProGlySerAsnLeuIleValAlaArgArgGly 487  
Db 1579 ACCAATGCTGCGCCCAAGCAAAAGATTAATCAACCTTCGCTTACAG---1622  
Qy 488 CysAsnProAlaProAlaProGlyCysSerProSerCysAla-----LeuThr 504  
Db 1633 ACAAAATGATATTTCTGATGTTCAAGATCTCACCAACATTAAGAGCATTAATATGTAAT 1692  
Qy 505 GlnLeuSerPheHisIleThrProThrAspSerLeuGluTrpHisGluAsnLeuGlyHis 524  
Db 1693 CAATGCTTTTCAACAAATCAGAAATGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1752  
Qy 525 GlySerPheThrIleIlePheArgGlySerArgArgGluValAlaAsp---GlyGluThr 543  
Db 1753 GGTACTTTTACAAAATTTTAAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1812  
Qy 544 HisAspSerGluValLeuLeuValIleMetAspSerArgHisArgAsnCysMetGluSer 563  
Db 1813 CACAAAG 1872

Qy 564 PheLeuGluAlaAlaSerLeuMetSerGlnValSerTrpHisIleuValLeuLeuHis 583  
Db 1873 TTTCTGAGCAGCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932  
Qy 584 GlyValCysMetAlaGly---AspSerIleMetValGlnGluPheValIleuValAla 602  
Db 1933 GGTGTCTGT 1992  
Qy 603 IleAspMetTrpLeuArgValArgGlyHisIleuValSerAlaSerTrpPheLeuGluVal 622  
Db 1993 CTGATACATTAACCTGAG 2052  
Qy 623 ThrGlyGlnLeuAlaIleThrAlaLeuAsnTrpIleuGluAspGlyGlyLeuProHisGlyAsn 642  
Db 2053 GCTAAGCAGTGTGATGAGGCGCCATGCTTTTACAGAAAGAAATCCCTTAATCATGGAAAT 2112  
Qy 643 ValSerAlaArgValLeuLeuValAlaArgGluGly-----AspGlyAsnProPro 660  
Db 2113 GTGTGTGCTAAATATATCTGCTTATTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2172  
Qy 661 PheIleLeuLeuSerAspProGlyValSerProThrValIleSerLeuGluMetLeuThr 680  
Db 2173 TTTATCAAACTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2232  
Qy 681 AspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700  
Db 2233 GAGAGAAATACATGGGTATCTCTGTAATGATGATGATGATGATGATGATGATGATGATGAT 2292  
Qy 701 AlaAspLysTrpGlyPheGlyAlaIleThrTrpGluValPheGlnAlaArgGlyProAlaHis 720  
Db 2293 ACAGACAGTAGTGAAGTGTGGAGACACCTGTGGAGATGTGCACTGAGAGAGATTAACCC 2352  
Qy 721 IleThrSerLeuGluProAlaValLysLeuLysLeuLysPheTrpGluAspGlnGlyLeuPro 740  
Db 2353 CTGAGT 2412  
Qy 741 AlaLeuLysTrpTrpGluLeuAlaGlyLeuIleThrGlnCysMetAlaIleTrpAspProGly 760  
Db 2413 GCACCCAAAGTGTGACAGAGATTAACCAACCTTAATTAATTAATTAATTAATTAATTAATTAAT 2472  
Qy 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTrp 780  
Db 2473 TTTGAGGCTGCTTTTCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2532  
Qy 781 GluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCysVal 799  
Db 2533 GAACTACTTAACAGAA---AATGACATGCTACCAACATGGAATAGAGTGGCTTGAAGGTTT 2589  
Qy 800 AlaGlyAlaGlnLeuTrpAlaCysGlnAspProAlaIlePheGlnGluArgHisIleuLys 819  
Db 2590 TCTGTGCTCT-----TTTGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2643  
Qy 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValIleLeuCysArgTrpArgAspPro 839  
Db 2644 TTTCTTACAGCAGCTTTGGCAAGGTAATCTTGGAGAGTGTGGAGATGTGGCGGTATGACCGG 2703  
Qy 840 LeuGlyAspAsnThrGlyProLeuValAlaValGlyGlnLeuGlnHisSerValProAsp 859  
Db 2704 CTGCAGAGACACACTGCGAGGT 2763  
Qy 860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuValAlaLeuHisSerAspPheIle 879  
Db 2764 CACCTCCGAGACTTTGAGAGGAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 2823  
Qy 880 ValLysTrpArgGlyValSerTrpGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899  
Db 2824 GTCAATGTAAGAGAGT 2883  
Qy 900 TyrLeuProSerGlyCysLeuArgAspLeuGlnGlnArgHisArg---GlyLeuHisThr 918  
Db 2884 TATTTAATCATATGAGAGTTTACGAGACTATCTCCAAAACATTAAGAGAGAGATGATATCAC 2943

```

QY 919 Asparagineutrophealatripginilecyslysglymetglutryleuglyalaarg 938
   :::::::::: ::::::::::::::::::::::::::::::::::::::::::::::::::
Db 2944 AAAAAATTTTCAATACATCTCAGATATGCAAGGCATGGAATATCTGTGTACAAA 3003
QY 939 ArgcysValHisargaspLeuAlaIaArgAsnIleuValGlsuerguAlaHisVal 958
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3004 AGGTATATCCACAGGGCCCTGGCAACAGATATGTGTGAAAAATGAAACAGGGTT 3063
QY 959 LysIleAlasphegilyLeuAlaIaLysLeuLeuProleuGlyLysAspTyrTyrValVal 978
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3064 AAATATAGAGACTTCGATTAACCAAGTCTGCCGAGCAAGAAATATCTACAAAGTA 3123
QY 979 ArggluProglyInsuSerProIlePheTyrAlaProgluSerIleuSerAspAsnIle 988
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3124 AAGAGCCAGGGGAAACCCCATATCTGTACCCACCTTAATCTTGACGGAGAGCAAG 3183
QY 999 PheSerArgInsuSerAspValTyrPheTyrValValIleuTyrGluLeuPheThrTyr 1018
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3184 TTTTCTGGCCCTCAGATGTGTGAGCTTTGAGACTGTTCATATGCAACTTTTCACATAC 3243
QY 1019 CysasplysSerCysSerProSerAlaGluPheLeuArgmetMetGlyProGluArgGlu 1038
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3244 ATCGAAGAGAGTAAAGTCACCCGTGGAATTTATGCAATGATGGCAATGATAAACAA 3303
QY 1039 GlyProProLeuGlyS---ArgLeuLeuGluLeuLeuAlaGluArgArgLeuProPro 1057
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3304 GGGCAATGATGTGTTCATTTGATGATAGCTACTGAAAGCAACGAGATGCCAAG 3363
QY 1058 ProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTTPAlaProGluPro 1077
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3364 CCAGAGAGATGCCCGCATGATGATTTATGTGATCAGACAGAGTCTGGAACACAAATGTG 3423
QY 1078 HisAspArgProAlaPheAlaThrLeuSer 1087
   ::::::::::::::::::::::::::::::::::::::::::::::::::
Db 3424 AGCCAGCGTCCCTCCTTCAGGAGACTTTTCG 3453

RESULT 9
AAD03607
ID AAD03607 standard; cDNA: 3629 BP.
XX AAD03607;
AC AAD03607;
DT 19-JUN-2001 (first entry)
XX
DE Murine (Janus kinase 2) Jak2 cDNA.
XX
XX Mouse; Janus kinase 2; Jak2; therapy: cytokine; tyrosine kinase;
KW antiproliferative; cytostatic; cell proliferative disorder;
KW cellular response; erythrocytosis; ss.
XX
OS Mus sp.
XX
XX Key Location/Qualifiers
FH 5'UTR 1..93
FT /tag= a
FT 94..3483
FT /tag= b
FT /product= "Murine (Janus kinase 2) Jak2 protein"
FT misc-feature 523
FT /tag= c
FT /note= "5' end of published partial Jak2 cDNA
FT sequence (Harpur et al.)"
FT /tag= d
FT replace (1089, T)
FT /tag= e
FT replace (1103, C)
FT /tag= f
FT replace (1111, G)
FT /tag= g
FT replace (1119, G)
FT /tag= h
FT replace (1122, C)
FT conflict

```

```

FT conflict
FT /tag= i
FT replace (1128, C)
FT /tag= j
FT replace (1131, G)
FT /tag= k
FT replace (1134, G)
FT /tag= l
FT replace (1137, C)
FT /tag= m
FT replace (1140, G)
FT /tag= n
FT replace (1143, G)
FT /tag= o
FT replace (1146, C)
FT /tag= p
FT replace (1188, T)
FT /tag= q
FT replace (1194, G)
FT /tag= r
FT replace (1230, G)
FT /tag= s
FT replace (1245, T)
FT /tag= t
FT replace (1260, T)
FT /tag= u
FT replace (1266, C)
FT /tag= v
FT replace (1272, T)
FT /tag= w
FT replace (1275, Q)
FT /tag= x
FT replace (1293, T)
FT /tag= y
FT replace (1305, T)
FT /tag= z
FT replace (1323, C)
FT /tag= aa
FT replace (1341, A)
FT /tag= ab
FT replace (1344, A)
FT /tag= ac
FT replace (1359, G)
FT /tag= ad
FT replace (1365, Q)
FT /tag= ae
FT replace (1368, T)
FT /tag= af
FT replace (1374, T)
FT /tag= ag
FT replace (1401, C)
FT /tag= ah
FT replace (1413, C)
FT /tag= ai
FT replace (1431, T)
FT /tag= aj
FT replace (1453, Q)
FT /tag= ak
FT replace (1476, G)
FT /tag= al
FT replace (1488, T)
FT /tag= am
FT replace (1511..1512, GT)
FT /tag= an
FT replace (1578, C)
FT /tag= ao
FT replace (1590, T)
FT /tag= ap
FT replace (1593, T)
FT /tag= aq
FT replace (1602, G)
FT /tag= ar
FT replace (1623, T)
FT /tag= as
FT conflict

```

```

FT conflict replace (1642, G)
FT /tag- at
FT conflict replace (1657, C)
FT /tag- au
FT conflict replace (1728, G)
FT /tag- av
FT conflict replace (1743, C)
FT /tag- aw
FT conflict replace (1755, C)
FT /tag- ax
FT conflict replace (1770, A)
FT /tag- ay
FT conflict replace (1809, G)
FT /tag- az
FT conflict replace (1816, G)
FT /tag- ba
FT conflict replace (1821, C)
FT /tag- bb
FT conflict replace (1857, A)
FT /tag- bc
FT conflict replace (1878, T)
FT /tag- bd
FT conflict replace (1935, A)
FT /tag- be
FT conflict replace (1938, A)
FT /tag- bf
FT conflict replace (1963, T)
FT /tag- bg
FT conflict replace (1974, G)
FT /tag- bh
FT conflict replace (2025, T)
FT /tag- bi
FT conflict replace (2055, G)
FT /tag- bj
FT conflict replace (2079, C)
FT /tag- bk
FT conflict replace (2082, C)
FT /tag- bl
FT misc_feature 2226
FT /tag- bm
FT /note- "Nucleotides corresponding to 7 amino acid insert"
FT conflict replace (2253, A)
FT /tag- bn
FT conflict replace (2259, G)
FT /tag- bo
FT conflict replace (2283, A)
FT /tag- bp
FT conflict replace (2285, C)
FT /tag- bq
FT conflict replace (2433, G)
FT /tag- br
FT conflict replace (3453, C)
FT /tag- bs
FT 3'UTR 3484..3629
FT /tag- bt
FT conflict replace (3579, C)
FT /tag- bu
FT conflict replace (3595, A)
FT /tag- bv
FT conflict replace (3599, A)
FT /tag- bw
FT conflict replace (3620, A)
FT /tag- bx
XX US6210654-B1.
XX 03-APR-2001.
XX 08-OCT-1997; 97US-0946994.
XX 18-JUN-1996; 96US-0665574.
XX 29-JUL-1993; 93US-0097997.

```

PA (SUDD-) ST JUDE CHILDREN'S HOSPITAL.  
 XX Ihle J, Witthuhn BA, Quelle FW, Silvennoinen O;  
 PI WPI; 2001-265367/27.  
 XX P-PSDB; AAE00352.  
 DR  
 XX  
 PT Modulating a biological response mediated by Jak kinase 2 activation to  
 PT a cytokine, useful for treating excessive proliferation of eukaryotic  
 PT cells, comprises inhibiting or enhancing tyrosine kinase activity of  
 PT Jak kinase in the cell -  
 XX  
 PS  
 XX Example 1: Fig 1, 100pp: English.  
 CC The present sequence is a cDNA encoding murine (Janus kinase 2) Jak2  
 CC tyrosine kinase. Jak2 sequence has a 600 amino acid long N-terminus that  
 CC lacks obvious SH2 (Src homology 2) and SH3 domains. Following this is a  
 CC kinase related domain (domain 2) and a carboxyl kinase domain (domain  
 CC 1). Jak kinases mediate cytokine activity through their tyrosine

Alignment Scores:  
 Pred. No.: 2,33e-190 Length: 3629  
 Score: 2619.50 Matches: 537  
 Percent Similarity: 65.50% Conservative: 190  
 Best Local Similarity: 48.38% Mismatches: 350  
 Query Match: 44.70% Indels: 33  
 DB: 22 Gaps: 15

US-09-397-967-16 (1-1099) x AAD03607 (1-3629)

```

QY 4 ProsergluGlutThrProleuileProgluArgSerCysleuSerSerSerGluAla 23
DB 145 CCTGACATCAGATGAGTATGTCCTGAGAGTCTGTAATCTGTAGACAGATAGACCA 204
QY 24 GluAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArgLeuSerPhe 43
DB 205 GTCTCTCAAGTATGATGTACCATTCCTGTGGCAAGCTGAGAGAGATGTGAGATT 264
QY 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIle 63
DB 265 CCAAGTGGAGATGATGTCTCCAGACAAATTTGTGGCGCTTCTTAACCTTGCTGTATT 324
QY 64 LeuProValTyrHisSerLeuPheAlaLeuAlaTyrGluAspPheSerCysTrpPhePro 83
DB 325 ACGCTGTGTATCATATATATGTTCGTTATAGAGTAAACGGAAGATGTGTATACCA 384
QY 84 ProSerHisIlePheCysIleGluAspValAspTrpGlnValLeuValTyrArgLeuArg 103
DB 385 CCCAATCATGTCTCCACATAGACGATCACCAGCATGACATGATCTACAGATTAAG 444
QY 104 PheTyrPheProAspTrpPhe-----GlyLeuGluTThrCysHisArgPheGlyLeuArg 121
DB 445 TTCTACTTCCCATTTGGTGTAGTGTAGTGGACAGACAGAACCTACAGATAGGAGTGC 504
QY 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141
DB 505 CTGGGGGTGAGAGCTGCTGTGATGACTTGTGCATGTCTTACTTTTCTGCAAGTGG 564
QY 142 ArgSerAspLeuValSerGlyArgLeuProValIleGlyLeuSerMetLysGlnGlnGly 161
DB 565 CGGCATGATTTGTTCCAGCGATGATAAAGTACCTGTACATCAATCAAGTGAAGAG 624
QY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlnGlnAlaGlnPro 181
DB 625 TGTCTTGGATGGCGGTGTAGACATGATGACATATAGCTAAGAGAAACAGCACTCA 684
QY 182 GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201
DB 685 CTGCGTGTCTATAACTGTCTACAGTAAAGACATCTTACCAAAAGTGGTGCAGCAAG 744
QY 202 IleGlnGlyGlnAspPheValThrArgArgArgIleArgArgThrValValLeuAlaLeu 221
DB 745 ATCCAAGACTATACATTTTAAACCCGGAAGCGAATCAGATTCAGATTTCGAGATTCT 804

```

QY 222 LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaTyrTyrIleLeuAsp 241  
DB 805 CAGCAATTCAGTCAATGTAAGCCAGCCAGCAAGCACTAAACCTTAAGTATCATATAAC 864  
QY 242 LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal----- 256  
DB 865 CTGCAAGCCCTGCAGCTCCTTCTACACAGAACAGCTTGAAGTAAAGAAATCTGCAGA 924  
QY 257 GlyLeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyTle 276  
DB 925 GGT---CCTTCAGGTGAGAGATTGTCACACCATTTAATACTGGGAACGGTGGAAAT 981  
QY 277 ProTyrPheSer-----AsnAspGluLeuPheGlnThr 287  
DB 982 CAGTGGTCAGAGGGGAACAATAAGGAAGTGAAGACCTGACAGAACAGAGAGCTACAGTA 1041  
QY 288 PheCysAspPheProGluIleValAlaSerIleAsnGlnAlaProArgValGlyPro 307  
DB 1042 TATGTGATTTCCCTGATATGATGTGATGCTAATTAAGCAAGCAACCAAG---GAATGC 1098  
QY 308 AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu 327  
DB 1099 TCAATGTAAGAGTGAATTTGACTGCTCATTAACAGAGTGTAAAGTTTGGAGATGAA 1158  
QY 328 PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347  
DB 1159 CTTAGCTCATTTAAAGAGAGCTGTGATGCTCATTTAATTAAGCGGTATTAACAGACTA 1218  
QY 348 IleCysAspSerArgHisTyrPheCysGlyGluValAlaProProAlaGluLeuGlu 367  
DB 1219 ACTGGGATGGGACCAATTTACTCTGCAAGAGGTGGCTCCCGAGCTGTGCTGAGAAC 1278  
QY 368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisIleLeuLysAla 387  
DB 1279 ATACACAGCACTGCCAGCGCCCAATATCAATGATTTTGCATTTAGCAAACTAAAGAG 1338  
QY 388 AlaGlySerLeuProGlyTyrThrIleLeuArgArgSerProGlnAspTyrAspSerPhe 407  
DB 1339 GCGGTAACAGACTGAGCATATGATGCTACAGTACAGCCCTAAGCACTCAACAAATAC 1398  
QY 408 LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyLysLeuIleArg 427  
DB 1399 TTTCTGACCTTGTGCTGTCGAGCAAAATGCTCAATTAATTAACCTGTTTATTAACG 1458  
QY 428 GluAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447  
DB 1459 AAGATAGAGATGAGATGAGATACCACTCAGCGGACTAAGAGACTTCAGTAACCTTAAG 1518  
QY 448 GluLeuLeuAlaAlaCysTTPAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 467  
DB 1519 GACCTTTTGAATTCCTACAGATGGAACCTGCGCTCAGACAGTATACATCTTCAGTTT 1578  
QY 468 ThrSerCysAlaProArgProLysGlnLysSerAsnLeuIleValAlaArgArgGly 487  
DB 1579 ACCAAATGCTGCCCCCAAGGCCAAAGATTAATCAAACTTCCTGCTTCAGA----- 1632  
QY 488 CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr 504  
DB 1633 ACAATGCTATTTCTGATGTTTCAGATCTCAACCACTTACAGAGCATTAATATGTGAAT 1692  
QY 505 GlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrHisGlnAsnLeuGlyHis 524  
DB 1693 CAATATGCTTTTCAAAACACAGAAATGAAGATTTAATTAATTAATGAAGCTTGGCCAA 1752  
QY 525 GlySerPheThrLysIlePheArgGlySerArgArgGluValAlaAsp---GlyGluThr 543  
DB 1753 GGTATCTTTTCAAAATTTTAAAGGTGTAAGAGAGAGATGAGATTTATGATCTG 1812  
QY 544 HisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSer 563  
DB 1813 CACAAACGAGAGTCTTTTGAAGTCTCTAGATTAAGCACATAGGAACCTATTCAGAGTCT 1872

---

QY 564 PheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuHis 583  
DB 1873 TTTCTGAGAGCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932  
QY 584 GlyValCysMetAlaGly---AspSerIleMetValGlnGluPheValTyrLeuGlyAla 602  
DB 1933 GGTGCTGCTGCTGCTGAGAGAGAGAACATTCCTGATCAAGATTTGTAATTTGGATCA 1992  
QY 603 IleAspMetTyrLeuArgLysArgGlyHisIleValSerAlaSerTyrIleLeuGlnVal 622  
DB 1993 CTGATACATACCTCGAAGAGAACAAAATTCCTAATATATATATGAAATCTGGAGTG 2052  
QY 623 ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnLysAspLysGlyLeuProHisLys 642  
DB 2053 GCTAAGACAGTTGGCATGGCCATGATTTCTAGAGAAAATCCCTTATTCAGGGAAT 2112  
QY 643 ValSerAlaArgLysValLeuLeuAlaArgGluGly-----AspGlyAsnProPro 660  
DB 2113 GTGTGCTGTAATAATATTCCTCTTATGAGAACAGACAGAGAGAGAGAGAGAGAGAGAG 2172  
QY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680  
DB 2173 TTTCACTCAACTAGTATGATCCCGGATTAAGCATTTACAGTTTACCGAAGAGATCTTCAG 2232  
QY 681 AspArgIleProTyrPheValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700  
DB 2233 GAGAGATACCATGAGTACCTCTGAATGCAATGAGATTCGTAATAATCTCAATCTGCA 2292  
QY 701 AlaAspLysTyrProGlyPheGlyAlaThrThrProGluValPheGlnArgGlyProAlaHis 720  
DB 2293 ACAGACAGAGTGAAGCTTGAGACCACTGTGGAGAGATCTGACAGTGAAGAGATTAAGCC 2352  
QY 721 IleThrSerLeuGluProAlaAlaLysLeuLysPheTyrGluAspGlnGlnLeuPro 740  
DB 2353 CTGAGTCTGCTGATTCCTCAAGAAAGCTGCACTTATAGAAATAGCATACACTTCTCT 2412  
QY 741 AlaLeuLysTyrPheGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760  
DB 2413 GCACCCAGAGTGAAGAGATTAAGCAACCTTAATAATTAATGCAAGATGAGACCGAGAT 2472  
QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlnLysLeuIleThrSerAspTyr 780  
DB 2473 TTCAGGCTGCTTTCAGAGCTGTGATCGGTATTTTAACAGCTGTTCATCCAGATTAAT 2532  
QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCysVal 799  
DB 2533 GAACCTACTAAGAGAA---AATGACATGCTACCAAAACATGAGAAATAGGCGCCCTAGGTT 2589  
QY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluAlaArgHisLeuLys 819  
DB 2590 TCTGTGCTCT---TTTGAAGACAGAGAGCCCTACACAGTTTGAAGAGAGACCTTGAAG 2643  
QY 820 TyrIleSerLeuLeuGlnGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPro 839  
DB 2644 TTTCTACAGCAGCTTGGCAAAAGTAACTTCGGAGTGTGAGAGATGTCCCGTATGAGCCG 2703  
QY 840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859  
DB 2704 CTGAGAGAACAACTGCGAGAGTGTGCTGTGAAGAAACCTCCAGCAGACACTGAAGAG 2763  
QY 860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879  
DB 2764 CACCTCGAGACTTTGAGAGGAGAGATGAGATCTCTGAATACTTTCAGCATGACATGAAC 2823  
QY 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899  
DB 2824 GTCAGATACAAAGAGAGTGTGATAGTGGCGGTGGCGGCAACCTTAAGATTAATTAAGAA 2883  
QY 900 TyrLeuProSerGlyCysLeuArgAspLeuGlnArgGlnArg---GlyLeuHisThr 918  
DB 2884 TATTTACATATGAGAGTTTACAGAGATCTCCAAAACATAAAGAGAGATGATGATCAG 2943  
QY 919 AspArgLeuLeuLeuPheAlaTTPGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg 938



```

Db 2944 AAAAACTTCTCAATCACATCTCAGATTCAGAGGCGATGGAATATATCTGTACAAAA 3003
      :::::
Qy 939 ArgCysValHisArgAspLeuAlaArgAsnIleLeuValIleuSerGluAlaHisVal 958
      ||| :::::
Db 3004 AGGTATATCCACAGGAGCCTGGCAACAGAACATATGTGTGAAATGAGAACAGCGTT 3063
Qy 959 LysIleAlaAspPheGlyLeuAlaLysLeuProLeuGlyLysAspTyrTyrValIle 978
      ||| :::::
Db 3064 AAATATGAGACTTCGGATTACCAAAAGCTTCCCGACAGCAAAATACTACAAAGTA 3123
Qy 979 ArgGluProGlyInserProIlePheTyrTyrAlaProGluSerLeuSerAspAsnIle 998
      :::::
Db 3124 AAGGACCCAGGGGAAAGCCCATATTCGTACGACCTGAAATCCTTTCAGGAGAGCAAG 3183
Qy 999 PheSerArgGluInserAspValTyrSerPheGlyValIleLeuTyrGluLeuPheTyr 1018
      ||| :::::
Db 3184 TTTTCTGTGGCCCTCAGATGTGTGAGACTTGGAGTGTCTATACCAACTTTTTCACATAC 3243
Qy 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetLeuGlyProGluArgIle 1038
      :::::
Db 3244 ATCGAGAACAGTAAAGTCCACCCCGTGAATTTATCGAATGTGGCAATGATAACAA 3303
Qy 1039 GlyProProLeuCys--ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPro 1057
      ||| :::::
Db 3304 GGGCAATATGATGTGTCTCATTGTAGAGCTACTAGACAGCAACGGAAGATTGCCAAG 3363
Qy 1058 ProProThrCysProThrGluValGlnIleuMetGlnLeuCysTyrAlaProGluPro 1077
      ||| :::::
Db 3364 CCAGAGAGATGCCAGATGAGATTATGTGATCATGACAGAGATGCTGGACAAACATGTG 3423
Qy 1078 HisAspArgProAlaPheAlaThrLeuSer 1087
      ||| :::::
Db 3424 ACCCAGCGCTCCTCTCTCAGGAGCCTTTCG 3453

RESULT 10
AAQ25307
ID AAQ25307 standard; DNA; 3473 BP.
XX
AC AAQ25307;
XX
DT 18-FEB-1999 (first entry)
XX
DE JAK2 encoding DNA.
XX
KW Phosphorylation; JAK1; JAK2; protein tyrosine kinase; human;
XX catalytic domain; SH2 domain; growth factor receptor; PTK; murine; ss.
XX
OS Mus musculus.
XX
FH Key location/Qualifiers
FT cds 1..2961
FT /*tag= a
FT /product= JAK2
FT
XX
XX MO9210519-A.
XX
XX 25-JUN-1992.
XX
XX 26-NOV-1991; 91WO-US08889.
XX
XX 28-NOV-1990; 90AU-0003594.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Harpur A, Wilks AF, Ziemlecki A;
XX
XX WPI; 1992-234591/28.
XX
XX P-PSDB; AAR25141.
XX
XX Novel protein tyrosine kinase mol. - comprises multiple catalytic
XX PT domains but no SH2 domain and is for phosphorylation of proteins
XX

```

```

PS Claim 10; Fig 8; 50pp; English.
XX
CC This sequence encodes the murine protein tyrosine kinase JAK2 (from
CC Janus kinase). Northern analysis of JAK2 expression in a mouse
CC demonstrated the presence of two mRNA transcripts (4.8 and 4.4 kb).
CC The levels of these transcripts alter with respect to one another in
CC different tissues. The kidney, spleen and lung appear to express
CC predominantly the larger form, whereas ovary, placenta, skeletal muscle
CC and all murine cell lines analysed express both forms at equal levels.
CC The difference in sizes may be due to differential polyadenylation
CC sites. Both JAK2 and JAK1 are examples of a new subfamily or class
CC of protein tyrosine kinase. These can be used in the phosphorylation
CC of proteins, incorporation of labels and in the design of analogues,
CC antagonists and agonists of JAK's.
XX
SQ Sequence 3473 BP; 1128 A; 677 C; 781 G; 887 T; 0 other;
XX
Alignment Scores:
Pred. No.: 8,49e-172 Length: 3473
Score: 2376.50 Matches: 489
Percent Similarity: 66.57% Conservative: 166
Best Local Similarity: 49.70% Mismatches: 298
Query Match: 40.55% Indels: 31
DB: Gaps: 14
XX
US-09-397-967-16 (1-1099) x AAQ25307 (1-3473)
Qy 128 IleuAspLeuHisValLeuGlnHisLeuPheAlaGlnHisArgSerAspLeuValSer 147
      :::::
Db 1 CTGCTTGATGACTTGTATCTTACCTCTTACCTCTTGTGTCGGCGCATGATTTGTCAC 60
Qy 148 GlyArgLeuProValGlyLeuSerMetLysGluGlnGlyLeuPheLeuSerLeuAlaVal 167
      ||| :::::
Db 61 GGATGGATAAAGTAACTACCTGTGATCACTGAACCTGAGAGAGTGTGGATGGCGCTG 120
Qy 168 LeuAspLeuAlaGlnAlaArgGluGlnAlaArgProGlyGluLeuLysTyr 187
      ||| :::::
Db 121 TTAGACATGATGAGAAATAGCTAAGAGAACAGCACTCCACGTGGCTGTACTAACTCT 180
Qy 188 ValSerTyrLysAlaCysLeuProProSerLeuArgAspValIleGlnGlyGlnAsnPhe 207
      ||| :::::
Db 181 GTCAGCTACAGACATCTTACCAAGTGGCTTCGACGGAAGATCCAAAGCTATCACATTT 240
Qy 208 ValThrArgArgArgIleArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeu 227
      :::::
Db 241 TTAAACCGGAAGCAATCAGGTACAGATTTCCGAGATTCATTCAGCAATTCAGTCAATGT 300
Qy 228 ProGlyArgProGlyAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisPro 247
      ||| :::::
Db 301 AAGCCACCTGCCAGAGAACTTAAAGTAACTTAAAGTAACTTAAAGTAACTTAAAGTAA 360
Qy 248 AlaAlaThrThrGluThrPheArgVal-----GlyLeuProGlyAlaGln 262
      ||| :::::
Db 361 GCCCTTACACAGAACGTTTGAAGTAAGAATCTGCAGAGAGT---CCTTCAGGTAG 417
Qy 263 GluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyTyrLeuProTyrSerSer----- 280
      ||| :::::
Db 418 GAGATTTTTCACACCATTAATAACTGGAAGACGTGGAATTCAGTGGTCAAGAGGAAA 477
Qy 281 -----AsnAspGluLeuPheGlnThrPheCysAspPheProGlu 293
      :::::
Db 478 CATAGGAAGATGAGACACTGACAGAAAGAGCTACAGTTATATGTGATTTCCCTGAT 537
Qy 294 IleValAspValSerIleAsnGlnIleProArgValGlyProAlaGlyGlnHisArgLeu 313
      ||| :::::
Db 538 ATTATGTATGTCAAGTTTAAAGCAAGCAACCAAG---GAATGCTCAAAATGAAAGTAAAT 594
Qy 314 ValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGlu 333
      ||| :::::
Db 595 GTAACGTGTCATGAACAGATGTAAGTTTGGACATAGAACTTACGTCATTAAGAAAGAA 654
Qy 334 AlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHis 353
      ||| :::::

```

```
Db 655 GCCTTGCTGATTCGTGATTAATTTGACGGGTATTTACAGACTAATCCGATGCGACCAAT 714
QY 354 TTTPhcCysLysGluValAlaProPArgLeuGluGluGluAlaAspValCysHis 373
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 715 TACCTTGCAAAAGAGTGGCTCCCGACAGCTGCTCGAAGAACTATACAGCAACTGCCAC 774
QY 374 GlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGly 393
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 775 GGCCCAATATCAATGATTTTGGCATTGACAAATTAAGAGCGGGGTACAGACTGGA 834
QY 394 ThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysVal 413
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 835 CTATATGCTGACAGTACGACCCCTTAAGACTTACAAATACCTTGTGACCTTGGCTGTT 894
QY 414 GlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAla 433
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 895 GACGAGAAATATGATTCATTTGAATTAACACTGTTGATTCGAAAGATGAGATGAGAA 954
QY 434 PheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgLysLeuLeuAlaCys 453
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 955 TACAACTTCAGCGGACTAATAGAACTTACAGTAACTTAAAGACTTTGAATGCTGAC 1014
QY 454 TrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaPro 473
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1015 CAGATGGAACCTGCGCTCAGACAGTATCATCTCCAGTTTACCAAACTGCTCCGCCCA 1074
QY 474 ArgProLysGluLysSerAsnLeuIleValAlaArgArgLysCysAsnProAlaProAla 493
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1075 AAGCCAAATATTAATCAAACTTCTGCTTCAGA-----ACAAATGGTATTTCGAT 1128
QY 494 ProGlySerProSerCysAla-----LeuThrGlnLeuSerPheHisThr 510
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1129 GTTCAGATTCACCAACATTAACAGAGCATTAATATGTAATCAAAATGGTTTCACAAA 1188
QY 511 IleProThrAspSerLeuGluTyrPheHisGlnLeuGlyHisGlySerPheThrLysIle 530
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1189 ATCAGATGAGATTAATTAATTAATGAAGATCTTGCCCAAGGTACTTTTACAAAAAT 1248
QY 531 PheArgGlySerArgArgGluValAlaAsp---GlyGluThrHisAspSerGlyValLeu 549
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1249 TTTAAAGGTGTAAGAGAGAGAGTTGAGATTTATGCTCACTGCGCAAAACGGAAGTCTT 1308
QY 550 LeuLysValMetAspSerArgHisArgAsnCysMetGlnSerPheLeuGluAlaIleSer 569
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1309 TTGAAGACCTAGATTAAGCAATAGCAACTATGACAGATCTTCTCGAAGCGCAAGC 1368
QY 570 LeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGly 589
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1369 ATGATGATTCAGCTTTTCACACAGCATTTGGTTTGAATTATGGTGTGCTGTGTGGA 1428
QY 590 ---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1429 GAGGAGAACATTCGTGTTCAAGCAATTTCTAAATTTGGATCAGCAATACCTGCTGAG 1488
QY 609 LysArgGlyHisLeuValSerAlaSerTyrPlyLeuGlnValThrLysGlnLeuAlaTyr 628
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1489 AAGAAACAAATATTCATTAATATTTATGAAACCTTGAGAGCGCTTAACACATGGCATAG 1548
QY 629 AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaGlyVal 648
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1549 GCCATGCAATTTCTAGAAAGAAATCCCTTATTCATGCGAATGTCTGTGCTAAATAATTC 1608
QY 649 LeuLeuAlaArgGluGlyLys-----AspGlyAsnProProPheIleLysLeuSerAsp 666
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1609 CTGCTTATCAAGAGAAAGACAGGAGAACGGGAAACCCACTTCATCAAACTTATGAT 1668
QY 667 ProGlyValSerProThrValLeuSerLeuGluMetLeuThrAspArgIleProThrVal 686
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1669 CTTGCAATTAGCATTAAGATTTCTACGAGACATTTTCAGAGAGAAATCCATCGGCTA 1728
QY 687 AlaProGlyCysLeuGluGluAlaGlnThrLeuCysLeuGluAlaAspLysTrpLysPhe 706
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1729 CCTCTGAATGATTTGAGATCTTAATAATCTCATCTGCGCAACAGACAGAGGAGCTTC 1788
QY 707 GlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeuPro 726
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1789 GGGACCACTCTGTGGAGATCTTGACAGTGGAGAGATTAAGCCCTTACAGCTCTGATTTCT 1848
QY 727 AlaLysLysLeuLysPheTyrGlnAspGlnGlnGlnLeuProAlaLeuLysTrpThrGlu 746
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1849 CAAGAGAGCTCAGTTCTATGAGATTAAGCATCAGCTTCTCGCACCCCAAGTGGACAG 1908
QY 747 LeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArg 766
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1909 TTAGCAAACTTAAATTAATGATGAGACTGATGAGCCAGATTTTCAGGCTGCTTCA 1968
QY 767 AlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspPro 786
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1969 GCTGTATCCGATCTTATACAGCTGTTTACTCCAGATTTAGAACTACTAACAGAA--- 2025
QY 787 ThrProGlyIleProSerProArg---AspGlyLeuCysValAlaGlyAlaGlnLeuTyr 805
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2026 AATGACATCTCAACCAACATGAGAAATAGGTGCCCTAAGGCTTTCTGTGCT---TTT 2079
QY 806 AlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLysTyrIleSerLeuGly 825
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2080 GAAGACAGGAGCCCTACACAGTTTGAAGAGACACTTGAATTTCTACAGCAGCTTGGC 2139
QY 826 LysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGly 845
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2140 AAGGTAACTTCGGGAGTGGAGATGTCCCTATGAGCCCTCAGAGCAACACTGGC 2199
QY 846 ProLeuValAlaValLysGlnLeuGlnHisSerValProAspGlnGlnArgAspPheGln 865
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2200 GAGGTGCTGCTGTAAGAAATCTCAGACACAGACTGAAGACACTCCGAGAACTTTGAG 2259
QY 866 ArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyVal 885
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2260 AGGAGATGAGATCTCGAATATCTTGACAGCATGACAAACATGCTCAAGTAAAGGAGT 2319
QY 886 SerTyrGlyProGlyArgGlnSerLeuAlaGlyLeuValMetGlnTyrLeuProSerGlyCys 905
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2320 TGTCTACAGTCCGGGTGGCGCACTAAGATTAATTAATGAAATTTTACCATATGCAAGT 2379
QY 906 LeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThrAspArgLeuLeuPhe 924
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2380 TTGAGACATATCTCCAAAACATAAGAAAGATGATGATCACAATAACTTCTTCATATC 2439
QY 925 AlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAsp 944
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2440 ACATCTCAGATATGCAAGGCGCATGGAATCTTGTATACAAAAGATATATCCACAGGAGC 2499
QY 945 LeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHisValLysIleAlaAspPheGly 964
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2500 CTGGCAACAGAGAACATATTGGTGAATAAGACAGGTTTAAATTAGAGACTTCGGA 2559
QY 965 LeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgLysProGlyGlnSer 984
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2560 TTAAACCAATCTTGGCCGCGAGACAAAGAAATACTACAAATTAAGAGACCGGAGGAAAGC 2619
QY 985 ProIlePheTrpTyrAlaProGlnSerLeuSerAspAsnIlePheSerArgGlnInsAsp 1004
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2620 CCCATATTCTGTGACGCACTCAATCTTGAAGAGAGCAAGTTTGTGCGCTCAGAT 2679
QY 1005 ValTrpSerPheGlyValIleLeuTyrGluLeuPheThrTyrCysAspLysSerCysSer 1024
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2680 GTGTGGAGCTTTGGAGTGTCTTATACGAACCTTTTCATATCATCGAAGAGATTAAGAT 2739
QY 1025 ProSerAlaGluPheLeuArgMetGlyProGlnArgGluGlyProProLeuCys--- 1043
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2740 CCACCCCGGAAATTTATGCGATGATTTGCAATGATTAACAGGCAAAATGATTTGCTTC 2799
QY 1044 ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProProProThrCysProThr 1063
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2800 CATTTGATAGACTTACGAGACAGCAACGGAAGATTTGCCAAGGCCAGAGATGCCAGAT 2859
```



QY 414 GlnThrProLeuGlyProAspTyrLysGlyCysLeuLeuLeuArgLysAspProSerGlyAla 433  
Db 895 GACGAGAAATGTTATTTGAATATAACACTGTTTGAATACAAAGATGAGATGGAG 954  
QY 434 PheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgLysLeuLeuAlaCys 453  
Db 955 TACACCTCGTGGACTAGAGAACCTTCAGTGTCTTAAAGACCTTTGGAATGGCTAC 1014  
QY 454 TrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaPro 473  
Db 1015 CAGATGGAACCTGGCGCTAGACAGATATCTTCAGTTCACCAATTCGCTCCCA 1074  
QY 474 ArgProLysGlnLysSerAsnLeuLeuValArgArgLysCysAsnProAlaProAla 493  
Db 1075 AACCCAAAGATAAATCAAACTCTTGTCTTCTGAGA-----ACAATGCTGTTTCTGAT 1128  
QY 494 ProGlyCysSerProSerCysAla-----LeuThrGlnLeuSerPheHisThr 510  
Db 1129 GTTCACCTCTCACCACTTACAGAGGCATTAATGTGATGCAATAGGTGTTTCAACAA 1188  
QY 511 IleProThrAspSerLeuGluTrpHisGluAsnLeuGlnLysGlySerPheThrLysIle 530  
Db 1189 ATCAGAAATGAATTTGATTTTAATGAAGCCTTGCCCAAGGCATTTTACAAATA 1248  
QY 531 PheArgGlySerArgArgLysValAlaAsp--GlyGluThrHisAspSerGluValLeu 549  
Db 1249 TTAAAGGTGTAAGAGAGAGAGAGATTGATGTGACGTGCAGCAACCGAAGTCTT 1308  
QY 550 LeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaIleSer 569  
Db 1309 TTGAAGGTCTAGATTAAGCAGCATAGAACTATTCAGAGCTTTCTTGAAGAGCAAGC 1368  
QY 570 LeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyAlaCysMetAlaGly 589  
Db 1369 ATGATGAGTACGCTTCTCAGCAGCACTTTGGTTTGAATTAAGAGATGTGCTGGCA 1428  
QY 590 ---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608  
Db 1429 GAGGAGAAATTTTGGTTCAMAGTTTGTAAATTTGGATCCTGATCATCATCCTGAG 1488  
QY 609 LysArgGlyHisLeuValSerAlaSerTyrLysLeuGlnValHisArgLysLeuAlaTyr 628  
Db 1489 AAGCAAAATTTCTTAATATATATATGAAACTTGGACGTGCGCAAGCAGTTGGCATGG 1548  
QY 629 AlaLeuAsnTyrLeuGlnAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648  
Db 1549 GCCATGCTCTCCTCGAAGAAATCCCTTATTCATGGAATGTGTGCTAAATATATC 1608  
QY 649 LeuLeuAlaArgLysGly-----AspGlyAsnProProPheLeuLysLeuSerAsp 666  
Db 1609 CTGCTTATCAGAGAAAG 1668  
QY 667 ProGlyValSerProThrValLeu-----SerLeuGlnMetLeu 679  
Db 1669 CCTGGCATTAAGCATTTACATCTTACCGAAGAGCATTTCTCCGTGTTTCCAGTTCTT 1728  
QY 680 ThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeu 699  
Db 1729 CAGGAGAGATACCATGAGGTGACACCTGAGTGATGAGATCTTAATAATTAACCTG 1788  
QY 700 GluAlaAspLysTyrGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAla 719  
Db 1789 GCAACAGCAAG 1848  
QY 720 HisIleThrSerLeuGluProAlaLysLysLeuLysPheThrGluAspGlnGlyGlnLeu 739  
Db 1849 CCCCTGAGTCTCTGATTTCTCAAGAAAGCTGAGTTCTATGAGATGAGATGAGCTT 1908  
QY 740 ProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759  
Db 1909 CCGGACCCCAAGTGGAGACAGTGGCAACCTTATTAATATTTGATGAGATGAGAGCA 1968

QY 760 GlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779  
Db 1969 GATTTCAGGCTGCTTTCACAGAGCTGATCATCCGATCTTTAAAGCCGTTTACTCCGAT 2028  
QY 780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProArg--AspGluLeuCys 798  
Db 2029 TATGACATCTAACAGAA---AATGACATGCTACCAAAACATGAAATAGCTGCTTACGG 2085  
QY 799 ValAlaGlyValGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluValArgHisLeu 818  
Db 2086 TTTTCTGTCT-----TTTGAAGACAGGAGACCCTTACACAGTTTGAAGAGAGACTTG 2139  
QY 819 LysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAsp 838  
Db 2140 AAGTTTCTACAGCAGCTTGGCAAAAGTAACCTGGGAGTGTGGAGATGTGCCCTATGAC 2199  
QY 839 ProLeuGlnAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPro 858  
Db 2200 CCGCTGAGAGCAAACTGGCGAGGTGTCTGTGTGAAGAACTCCAGCAGCAGCTGAA 2259  
QY 859 AspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPhe 878  
Db 2260 GACGACCTCCGAGACTTTTGAGAGGAGATCGAGATCTGAAATCTTGCAGCATGCAAC 2319  
QY 879 IleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMet 898  
Db 2320 ATCGTCAGTACAGAGAGAGTGTCTACAGTGGCGGCGGCAACCTAGATTAATATG 2379  
QY 899 GluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg--GlyLeuHis 917  
Db 2380 GAATATTATACCATATGGAAGTATTCAGAGACTATCTCCAAAACATGAAGAGAGATAT 2439  
QY 918 ThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAla 937  
Db 2440 CACAAAATCTTCTTCAATACATCATCTCAGATATGCAAGGCAATGGAATATCTTGATCA 2499  
QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957  
Db 2500 AAAAGGTATATCCACAGGCGCTGCGCAACAGAGAACTATTTGGTGGAATGAGAACAG 2559  
QY 958 ValLysIleLeuAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVal 977  
Db 2560 GTTAAATATAGAGACTTTCGATTAACCAAGTCTTCCCGAGAGCAAAAGATCTTCAAA 2619  
QY 978 ValArgGluProGlyLysIleSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsn 997  
Db 2620 GTAAG 2679  
QY 998 IlePheSerArgGlnSerAspValIleTrpSerPheGlyValAlaLeuTyrGluLeuPheThr 1017  
Db 2680 AAGTTTCTGTGGCTCAGATGTGTGAGCTTGGAGTGTCTTATACAACTTTTTCACA 2739  
QY 1018 TyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetGlyProGluArg 1037  
Db 2740 TACATGAGAGAGAGTAAAGTCCACCCGTGGAATTTATGCGAATGATTTGCGATGATTA 2799  
QY 1038 GlnGlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGlnLysArgLeuPro 1056  
Db 2800 CAAGGCAATATGATTTGTTCATTTGATATAGAGCTACTGAAAGCAAGAGAGATTTGCCA 2859  
QY 1057 ProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGlu 1076  
Db 2860 AGGCCAG 2919  
QY 1077 ProHisAspArgProAlaPheAlaThrLeuSer 1087  
Db 2920 GTGAGCAGAGCTCTCTTTCAGGAGCTTTCC 2952  
RESULT 12  
ABK72331  
ID ABK72331 standard; DNA; 4191 BP.  
XX  
AC ABK72331;

XX 30-JUL-2002 (first entry)  
 XX DNA encoding lymphoma associated mouse JAK1 protein.  
 DE  
 XX Lymphoma associated protein; PIK3R1, GNMS; XI-(alpha-s-); NESP55; JAK1;  
 KW Neurogranin; Nf12; Lymphoma; DNA vaccine; cytotoxic T-cell;  
 KW animal model; Leukemia; ds.  
 XX Mus sp.  
 OS  
 XX WO200224867-A2.  
 PN  
 PD 28-MAR-2002.  
 XX  
 XX 24-SEP-2001; 2001US-0529798.  
 PE  
 XX 22-SEP-2000; 2000US-0668644.  
 PR 13-JUL-2001; 2001US-0905390.  
 PR 13-JUL-2001; 2001US-0905491.  
 XX  
 XX (UYAA-) UNIV AARHUS.  
 PA  
 XX Pedersen FS, Sorensen AB, Hernandez JM, Nielsen AA, Moving HO;  
 PI  
 XX WPI: 2002-416279/44.  
 DR  
 XX P-PSDB: ABG60307.  
 DR  
 XX  
 PT Novel recombinant lymphoma associated protein (LAP) such as PIK3R1,  
 PT GNMS, JAK1, Neurogranin, Nf12 proteins, useful for identifying  
 PT inhibitors of LAP activity that are used for treating lymphoma -  
 XX  
 PS Claim 1: Page 148; 160pp; English.  
 XX  
 XX The invention describes a recombinant protein (I) from a lymphoma  
 CC associated protein (LAP) sequence of PIK3R1, GNMS (including  
 CC XI-(alpha-s-), and NESP55), JAK1, Neurogranin or Nf12 proteins. (I) is  
 CC useful for screening for a bioactive agent capable of binding to an LA  
 CC protein (LAP) which is encoded by a polynucleotide (II) and (II) is  
 CC useful for evaluating the effect of a candidate lymphoma drug in a  
 CC patient. (I) and (II) is also useful for diagnosing lymphoma involving  
 CC determining the expression of one or more (II), or (I) encoded by (II)  
 CC in a first tissue type of a first individual and comparing the expression  
 CC of the gene(s) from a second normal tissue type from the first individual  
 CC or a second unaffected individual, where a difference in the expression  
 CC indicates that the first individual has lymphoma. (II) is useful in the  
 CC diagnosis and treatment of lymphoma and leukaemia. A host cell  
 CC comprising (II) is useful for screening drug candidates which involves  
 CC providing the host cell that expresses LA gene, adding a drug candidate  
 CC to the cell, and determining the effect of the drug candidate on the  
 CC expression of the LA gene. LA genes are also administered as DNA  
 CC vaccine, such that expression of the polypeptide encoded by the DNA  
 CC vaccine, cytotoxic T-cells and antibodies are induced which recognise and  
 CC destroy or eliminate cells expressing LA proteins. The genes are also for  
 CC generating animal models of lymphoma. This sequence represents a lymphoma  
 CC associated polynucleotide described in the invention.  
 XX  
 SO Sequence 4191 BP; 1214 A; 982 C; 1024 G; 971 T; 0 other;

Alignment Scores:  
 Pred. No.: 6,88e-136 Length: 4191  
 Score: 1908.50 Matches: 437  
 Percent Similarity: 54.52% Conservative: 184  
 Best Local Similarity: 38.37% Mismatches: 383  
 Query Match: 32.57% Indels: 135  
 DB: 24 Gaps: 22

US-09-397-967-16 (1-1099) x ABK72331 (1-4191)

QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuGlyValAlaGlnAlaLysAla 60  
 DB 408 CTCGCCCTGGGCGAGAGATATACAGCCGAGAGAGCTGTGATCATCGGCGCCGCGAGAG 467

QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
 DB 468 TGCAGTATCTCTCTCTGTGTCACAACTTCGCGCTAGCATGAGACCAAGCTC 527  
 QY 81 TrpPheProPheSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
 DB 528 TGGTACGGCTCCCAACCGAATCATCACTGTGATGACAAACGTCCTCCGCTCCATGC 587  
 QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeu-----GluThrCysHis 116  
 DB 588 CGCATGAGGTTCTACTTTCACCACTGGCCAGCAACCAATGACACGAAACAGTCTGTATG 647  
 QY 117 ArgPheGlyLeuArgLysAspLeu-----ThrSer 126  
 DB 648 CGACATTTCTCCAAAGAGACAGAAAACGGCTGTGAGAGAAAAGGTTCCAGACAGAAC 707  
 QY 127 AlaIleLeuAspLeuHisValLeuGluHisIlePheAlaGlnHisArgSerIleVal 146  
 DB 708 CCACCTCTGTGATGCCAGTTCCTGAGTATCTGTTCACAGGACAGATATGATTTGATC 767  
 QY 147 SerGlyArgLeuProValGlyLeuSerMetLysGluGln-----Gly 160  
 DB 768 AATGCTGTGCTCCCATTTGGGAGCCCAAGAGAGAGACGAGACATGATATTGAAAT 827  
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArg 180  
 DB 828 GAGTGGCTGGGAGATGGCGGTCTCCGCTCCATCTCCCATGATGATGAAGATGACG 887  
 QY 181 ProGlyIleLeuLeuLysThrValSerTyrTyrAlaCysLeuProPheSerLeuArgAsp 200  
 DB 888 TTGCGGAACTTCCCAACACATCACTCAAGCATATATTCACAGAAACATTGAATAA 947  
 QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIle----- 213  
 DB 948 TCCATTCAGACAGAGAGACCTTCTTACAGAGATGCGAATAAATGTTTCAAGATTTC 1007  
 QY 214 -----ArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeuPro 228  
 DB 1008 TTGAAGAAATTAACAAACAGACATC-----TGTGAC 1040  
 QY 229 GlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGlu-----ArgLeuHis 246  
 DB 1041 AGCAGTGTCACTGAGCTGAAGTGAAGTAACTCTGCTACCTTGGAACCTTCACTTGACA 1100  
 QY 247 ProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGluLeuProGly 266  
 DB 1101 AAACATTATGAGCTGAATATTTGTGACTGTATGCTACTGATTCATCAGAAATGAA 1160  
 QY 267 LeuLeuArg-----ValAlaGly 272  
 DB 1161 TTGAGTCTGATGCTCATTCGAAATGACAGTGCATGTTCTGTATGAGGTGATGTGATG 1220  
 QY 273 AspAsnGlyIleProIlePheSerAsnAsp----- 282  
 DB 1221 AATCTCGGATTCAGTGGCGGAGAAACCAAAATGTTTCTCTGTGAAAGGAAAAAAT 1280  
 QY 282 ----- 282  
 DB 1281 AAACGTAGCGGAAAAAAGCTGAAATATTAACACAGAGATGATGAGAAACAA 1340  
 QY 283 -----GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsn 300  
 DB 1341 CTCGGGAGAGAGTGAACATTTTCTCTTTCCTCGAATACCCACATGTAATGAAG 1400  
 QY 301 GlnAlaProArgValGlyProAlaGlyGlnHisArgLeuValThrValThrArgMetAsp 320  
 DB 1401 GAGTCT-----GTGCTCAGCATTTAAACAAACAGGAC 1430  
 QY 321 GlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLys 340  
 DB 1431 AACAAACATGAGAACTCAAGCTCTTCTCGAAGAAAGCTTGTCTTTGTGCTCCG 1490  
 QY 341 ValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAla 360

```

Db 1491 GTGATGGCTACTTCGGCTCAGTCAGATGACACCACTACCTCTGACTGATGTGCT 1550
QY 361 ProProArgLeuLeuGluGluGluAlaAlaValCysHisGlyProIleThrLeuAspHe 380
Db 1551 CCCCACAGTATGTCACCAATATACAGAAAGGCTGCCAGGCTCCATGTGCAGAAATAT 1610
QY 381 AlaIleHisLeuLeuAlaAlaGlySerLeuProGlyThrIleLeuArgArgSer 400
Db 1611 GCCATCATATAGCTGGCGCAGGAAGGAGTGAAGGGGATGTACGTCTGAGGTGAGC 1670
QY 401 ProGlnAspTyrAspSerPheLeuLeuThr---AlaCysValGlnThrPro-----Leu 417
Db 1671 TGCACCGACTTGGACAAATCTTATGACTGTCACTGCTTGAAGAACTGTAGAGTATG 1730
QY 418 Gly-----ProAspTyrLeuGlySerLeuLeuArgGlnAspProSerGlyAlaAlaPheSer 435
Db 1731 GGTGGCCAGAAAGCAGTTCACAACTTTCAGATT---GAGGTACAGAAAGGCCCTACAGC 1787
QY 436 LeuValGlyLeuSerGlnProHisArgSerLeuArgGlnLeuLeuAlaAlaCysTyrAsn 455
Db 1788 CTGCATGGCTGTATGACACCACTTCCAGCTTGGACACTCATGAAACCACTCAAGAG 1847
QY 456 SerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysGlyAlaProArgPro 475
Db 1848 CAGATCTGCGCCAGCAGACATATAGCTTTGTGTGAACAGCATCTGTCACTCAAGCCT 1907
QY 476 LysGlnLysSerAsnLeuLeuValValArgArgGlyCysAsnProAlaProAlaProGly 495
Db 1908 CGAATAATCTCATCTGCTGCTGAGCCACTAG-----AAAGCCAGAG 1952
QY 496 CysSerProSerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSer 515
Db 1953 TGGCAGCT---GTCCTACTCATGAGCCAGCTTGTGAGCTTGTGAGCTTAAAGAAAT 2009
QY 516 LeuLeuThrPheHisGlnAsnLeuGlyHisGlySerPheThrIlePheArgGlySerArg 535
Db 2010 ATTATACAGGTGACACCTTGGCAGAGGACAAAGAACATATATCTTGGGAGCCCTG 2069
QY 536 ArgGlnValValAsp-----GlyGlnThrHisAspSerGlnValLeuLeuLys 551
Db 2070 CTGACATCAAGAGATGAGGAAGAAATGCTCAAGAGAAAGAAATATATCTCTCAA 2129
QY 552 ValMetAspSerArgHisArgAsnCysMetGlnSerPheLeuGlnAlaAlaSerLeuMet 571
Db 2130 GTCTTACAGCCACACCGGAGCATCTCTGCGCTTCTTGTGAGGCTGATGATG 2189
QY 572 SerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGly---Asp 590
Db 2190 AGCAGAGTTTCCCAACAAATATAGTGTACTTACGCGGTGTGTCCGAGATGTGAA 2249
QY 591 SerIleMetValGlnGluPheValTyrLeuGlnValAlaIleAspMetTyrLeuArgArg 610
Db 2250 AATATCATGTGGGAAGATTTGTGGAGGGGCGCGCTTGATCTCTTCATCACCAGGAAA 2309
QY 611 GlyHisLeuValSerAlaSerTyrPlyLeuGlnValThrLysGlnLeuAlaTyrAlaLeu 630
Db 2310 AGTCATCGCTTACTACCCCTGGAAGTCAAGGTTGCCAAACAGCTGGCCAGTGGCTG 2369
QY 631 AsnTyrLeuGlnAspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeu 650
Db 2370 AGTTACTTGAAGATTAAGACCTGTTGATGAAATGTGTGCACATAAAGACCTCTCTG 2429
QY 651 AlaArgGlnGlyLysArgLysAsn---ProPheIleLysLeuSerAspProGlyVal 669
Db 2430 GCCCGTAGGCGCATGTGACAGTATGCGCCCTTCATCAAGCTTATGAGCCCTGAGATC 2489
QY 670 SerProThrValLeuSerLeuGlnMetLeuThrAspArgIleProTyrValAlaProGln 689
Db 2490 CCAATCTCTGTGTCAGCAGCAAGATGATAGAGCAATCCCTGAGTGCCTGAG 2549
QY 690 CysLeuGlnGlnAlaGlnThrLeuCysLeuGlnAlaAspLysTyrPlyPheGlyAlaThr 709
|||||
Db 2550 TGTGTGAAGACTCCAGAACCTGATGTGGCTGCTGCAAGATGAGACTTTGAACACAG 2609
QY 710 ThrTyrGlnValPheGlnArgGlyProAlaHisIleThrSerLeuGlnProAlaLysLys 729
Db 2610 CTCTGGAAATCTGCTACACAGGAGATTTCTCTCAAGACAAAGACCTTCATTGAGAAA 2669
QY 730 LeuLysPheTyrGlnAspGlnGlyGlnLeuProAlaLeuLysTyrThrGlnLeuAlaGly 749
Db 2670 GAGAGGTTTATGAAGCCGTCGACAGGCTGTGATCTCATCTTGCAGAGGACTAGCTGAC 2729
QY 750 LeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeu 769
Db 2730 CTCAATGACTGCTGCTGATGAACATATGACCCCAACAGAAACCTTCTTCCGAGCATCATG 2789
QY 770 ArgAspLeuAsnGlyLeuIleThrSerAspTyrGlnLeuLeuSerAspProThrProGly 789
Db 2790 AGGACATTTAACAGCTGAGAGACAGAAATCCAGATGTTGTTAGAA-----GAC 2837
QY 790 IleProSerProArgAspGlnLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAsp 809
Db 2838 ---AAGCAGCCCAACACAGAGTG-----GAC 2861
QY 810 ProAlaIlePheGlnGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPhe 829
Db 2862 CCCACTCACTTGAAGAACCGGCTCTGTAAGAGATTTGTGTGACTTGGAGAGGTCACCTT 2921
QY 830 GlySerValGlnLeuCysArgTyrAspProLeuGlnLysAsnThrGlnProLeuValAla 849
Db 2922 GCGAAGGTGAGCTCTGAGATATGATCTTGAAGGACACACAGGAGGAGCTAGCT 2981
QY 850 ValLysGlnLeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGlnIle 868
Db 2982 GTCAAGCTCCCTGAGAGCTGAGAGTGAGGTAAACCAATAGCTGATCTGAAAGAGAGATA 3041
QY 869 GlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGly 888
Db 3042 GAGATCTTACGGAACCTCTCATGATGAGAACATTTGTGAAGTACAAAGAAATGTCATGAA 3101
QY 889 ProGlnArgGlnSerLeuArgLeuValMetGlnTyrLeuProSerGlyCysLeuArgAsp 908
Db 3102 GACGAGGCAATGTGTATCAACGCTCATGATGAGATTTGCTGCGTGGAGAGCTTAAAGAG 3161
QY 909 LeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuLeuPheAlaTyrGln 927
Db 3162 TATCTGCCAAAGAAATAGAACAAATCAACCTCAACAGCAGCTAAATATGCGCATCAG 3221
QY 928 IleCysLysGlnMetGlnTyrLeuGlnValAlaArgArgCysValHisArgAspLeuAla 947
Db 3222 ATTTGTAAAGGAGATGACTACTTGGTTCTGCAATATGCTTACCGGAGACTTAAAGACA 3281
QY 948 ArgAsnIleLeuValGlnSerGlnAlaHisValLysIleAlaAspPheGlyLeuAlaLys 967
Db 3282 AGAATATGCTCTGTGTGAGAGTGCATCAACATGAGTGAAGTCCGAGACTTTGGTTAACCAA 3341
QY 968 LeuLeuProLeuGlnLysAspTyrTyrValValArgGlnProGlyGlnSerProIlePhe 987
Db 3342 GCAATTTAAACCGATTAAGAGTACTACACAGTCAAGAGCAGCCGAGACGCCAGCTGTC 3401
QY 988 TyrTyrLarProGlnSerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrSer 1007
Db 3402 TGTGTACCTCCGGAATTTTATATCCAGTGAATTTTATATGCGCTGTGAGTGTGCT 3461
QY 1008 PheGlyValValLeuTyrGlnLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027
Db 3462 TTTGAGATGACACAGCAGAGCTCTCATCTTACGTGATCACTTATAGTCCATGAGCC 3521
QY 1028 GlnPheLeuArgMetGlyProGlnArgGlnLysProProLeuCysArgLeuLeuGln 1047
Db 3522 TTGTCTTGAAGAAATAGAGGCCCAACTCATAGGCCACATGACAGTACAGCGATTGTGAG 3581
QY 1048 LeuLeuAlaGlnGlyArgArgLeuProProProProThrThrCysProThrGlnValGlnGln 1067
Db 3582 ACTCTGAAGAGAGGACGCTGCTGATGTCACACCAACTGCTCTGATGAGGTTATCAG 3641
```

Qy	36	glyproglinrgrleuserPherPhegIyaSPrrYtLeuAlaGluaspLeucyVal	55
Db	106	-----CTCCGGCTGGGACGTAGAGGTTCACGACGAGAGAACTGTGCATTC	150
Qy	56	ArgAlaIaIaLysIaIaCysglYIleleuProValTYrHisSerLeuPheAlaIeuaIaThr	75
Db	151	AGGGCTCCACAGCAGCATTCGCGTATCTCTCTCTTTGTGTGCACACCTTTTGCCCTGTATGAC	210
Qy	76	GluAsPheSerCysTrpPheProPheSerHisIlePheCysIleGluAsPheIaSPrrY	95
Db	211	GAAGAACCCAAAGCTCGGTATGATCTCCAAATCCACCATCACCCGTTGATGACAGATGTC	270
Qy	96	GlnValIeuaValYrArgLeuArgPheTYrPheProAspTrpPheglYLeu	112
Db	271	CTCCGGCTCCACAGCAGCATTCGCGTATCTCTCTTTGTGTGCACACCTTTTGCCCTGTATGAC	330
Qy	113	---GluThrCysHisArPheGlyLeuArgIysAspleu	124
Db	331	GAGCAGTCACTGGTGGCTATCTCCAAAGAAAGAAATGGCTACGAGAAAAAAG	390
Qy	125	-----ThSerAlaIleLeuAspLeuHisValIleuGluHisLeuPheAlaGlnHis	141
Db	391	ATTCCAGATGCAACCCCTCTCTTATATGCCAGCTCACGTGAGATCTGTGGTCTCAGGA	450
Qy	142	ArgSerAspleuValSerGlyArgLeuProValIleLeuSerMetIysGluGln	159
Db	451	CAGTATGATTTGGTGAATGAGCTGCTGCTCTTATGAGAACCCCAAGACCAGACAGATGGA	510
Qy	160	-----GlyGluPheLeuSerIeuaIaValIeuaSPleuaIaGlnMetAlaArg	175
Db	511	CATGATATTGACAGACAGGTCTAGAGTGGCTCTCTCGGCACTCACACTATGGCATG	570
Qy	176	GluGlnAlaGlnArgProGlyGluLeuLeuIysThrValSerTYrIysAlaCysLeuPro	195
Db	571	ATGAGAGAGATCCAGTGCACAGACTGCCCAAGAGACATCAAGCTTCAAGCAGCATATTCCA	630
Qy	196	ProSerLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgIle	213
Db	631	GAAGACTTGAAATTAAGTCATCAGACAGAGAACCTTCTCACAGAGATCGCGATTAATAAT	690
Qy	214	-----ArgArgThrValIaIeuaIleuLeuLeuPro	223
Db	691	GTTTTCAGAGATTCCTTAAGAGATTTTAACACAAAGCCATTGTGCACACACCGTG	747
Qy	224	CysGlyArgLeuProGlyArgProTYrAlaIeuaMetAlaIysTYrIleAspLeuGln	243
Db	748	-----TCCACCCAGTACCTGAAAGGTGAATTAATCTGGCTACCTTGGAA	789
Qy	244	ArgLeuHisProAlaIaIaThrThrGluThrPheArgValIaIeuleuProGlyIaGlnGlu	263
Db	790	ACTTTGACAAAAACATTAACGGTGCCTGAATATT	822
Qy	264	GluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrp	282
Db	823	GAGACTTCACATGTTACGATTTCTATCAGAAATAGAGATGAATGTGTTTCATTCGATGAC	882
Qy	282	-----	282
Db	883	GGTGGAACGTTCTCTACTACGAAGATGATGGTACTGGGAATCTTGGAAATCCAGTGAGAG	942
Qy	282	-----	282
Db	943	CATAAACCAATGTGTTCTGTTGAAAGAGAAATAAATTAACCTGAAGCGGAAAAACTG	1007
Qy	283	-----GluLeuPheGlnThrPhe	288
Db	1003	GAATAATTAAGACAGAGATGAGGAGAAAAACMAAGATCCGGGAGAGAGTGAACAATTTT	1062
Qy	289	CysAsPheProGluIleValIaSPrrIleAsnGlnAlaIaProArgValIaIeProAla	308
Db	1063	TCATTTCCCTGAATCACTCACTATTGTATTAATTAAGAGATCT	1104



```

OY 309 G1yG1uH1s1rG1eUvAl1ThrVAl1ThrArGMe1AsP0Ly1H1s1I1eUg1uAl1aG1uPhe 328
      :::::::::::::: :::: :::: :::: ::::
Db 1105 -----GTGCTCACATTAACAGACAGACAAAGAAAGAAATGGAAGCTC 1152
OY 329 ProG1yLeuProG1uAl1aLeuSerPheVal1Al1eUvAl1sP0Ly1rPheArG1u1le 348
      :::::::::::::: :::: :::: :::: ::::
Db 1153 TCCTCCACAGAGAGGCGCTTGTCTTGTCTGCTGATGAGTGGCTCTCCGCTCACA 1212
OY 349 CysAsPserArG1H1s1rPheCysLyS1uVAl1Al1aProProArG1eUleUg1uG1u 368
      :::: :::: :::: ::::
Db 1213 GCAGATGCCATCATTTACCTCTGACCGAGTGCGCCCCCGTTGATGCTCCACAACTATA 1272
OY 369 AlAAsPvAlCysH1s1rG1Pro1leTh1LeuAsPheAl1a1eH1s1LyS1eUyAl1a1a 388
      :::: :::: :::: ::::
Db 1273 CAGAAATGGCTGTCAATGCTCCATCTGACAAATACGCCATCAATAAATTCGGCAAGAA 1332
OY 389 GlySerLeuProG1yTh1rTh1LeuArG1rSerProG1uAsP1rAsPserPheLeu 408
      :::: :::: :::: ::::
Db 1333 GGAAGCCGAGAGGAGTGTACTGTCTGAGTGAGCTGACCGCCTTGACACATCTCTC 1392
OY 409 LeuThr---AlAcysValG1nThrPro-----LeuG1yProAsP1rLyS 422
      :::: :::: :::: ::::
Db 1393 ATGACCTGCACCTGCTTGAAGAAGCTGACAGCGAGGCTGCCCAAGACAGTTCAG 1452
OY 423 GlyCysLeu1leArG1nAsPProSerG1yAl1aPheSerLeuVAl1G1yLeuSerG1nPro 442
      :::: :::: :::: ::::
Db 1453 AACTTTCAGATC---GAGGTCCAGAGAGGCGCTACACATCTGCAGCGGTTCGACCGCAC 1509
OY 443 H1sArGserLeuArG1uLeuAl1a1a1aCys1rPAsnSerG1yLeuArG1vAl1AsP1y 462
      :::: :::: :::: ::::
Db 1510 TTCCCCACCTTGCGAGACCTCATGACCACTCAAGAAAGACAGATCTCTGCGACGATMAC 1569
OY 463 AlAAlaLeuTh1rLeuTh1rSerCysAlaProArG1r0LyS1uLyS1eUyAl1eU 482
      :::: :::: :::: ::::
Db 1570 ATCAGCTTCACTGCTAAACGCTGCTCCAGCCCAAGCCCAAGAAATCTCAACCTGCTG 1629
OY 483 Val1ValArGArG1yCysAsnProAlaProAlaProG1yCysSerProSerCysAla 502
      :::: :::: :::: ::::
Db 1630 GTGGCTACTAAG-----AAAGCCAGAGTGGCACCC---GTCTACCCC 1671
OY 503 LeuThrG1nLeuSerPheH1s1rTh1leProTh1rAsPserLeuG1uTh1rH1s1uAsnLeu 522
      :::: :::: :::: ::::
Db 1672 ATGAGCAGCACTGATTCGATCGATCCCAAGAAAGATCTGTCAGGCGGACGACCTT 1731
OY 523 G1yH1s1rG1ySerPheTh1rLyS1lePheArG1ySerArG1uVAl1AsP1yG1u 542
      :::: :::: :::: ::::
Db 1732 GGGAGAGCGACAGACACACATCTATCTGGGACCCCTGATGATTACAGATACAGAA 1791
OY 543 ThrH1sAsPserG1u-----ValLeuLeuLyS1ValMetAsPserArG1H1sArG 558
      :::: :::: :::: ::::
Db 1792 GGAACCTTCGAAAGAAAGAAAGATTAAGTCTCAAGTCTTACGCCACGACCAAG 1851
OY 559 AsnCysMetG1uSerPheLeuG1uAl1a1a1aSerLeuMetSerG1nVAl1Ser1rProH1s 578
      :::: :::: :::: ::::
Db 1852 GATATTTCCCTGCTCTTCTTCGAGCGAGCCACATGATGATGAGCTCTCCCAACAAAC 1911
OY 579 LeuValLeuLeuH1s1rG1yAl1CysMetAlaG1y---AsPser1leMetAl1s1nG1uPhe 597
      :::: :::: :::: ::::
Db 1912 ATCGTGTACTCTATGAGCGCTCTGTCTCCGAGCTGGAGAAATATATCTGTGGAAGATT 1971
OY 598 Val1rLeuG1yAl1a1eAsPMet1rLeuArG1yLyS1ArG1yH1s1eUvAl1SerAl1aSer 617
      :::: :::: :::: ::::
Db 1972 TGGAGAGGGGCTCTCTGATCTCTTCATGACCGGAAAGATGATCTCTTACACACCA 2031
OY 618 TrpLyS1eUg1nVal1Th1rLyS1nLeuAl1a1rAl1aLeuAsnTrpLeuG1uAsP1yS1y 637
      :::: :::: :::: ::::
Db 2032 TGGAAATTCAAAGTGGCCAAACAGCTGGCCAGTGGCTTGAGCTTGGAGGATTAAAGC 2091
OY 638 LeuProH1s1rG1yAsnVal1SerAlaArG1yVal1LeuLeuAlaArG1uG1yG1yAsP1y 657
      :::: :::: :::: ::::
Db 2092 CTGTGCTCATGAAATGTGTACTATAAAACCTCTCTGCGCCGTGAGGAATCGACAGT 2151
OY 658 Asn---ProProH1e1leUyS1eUserAsPProG1yVAl1SerProTh1rVAl1eUserLeu 676
      :::: :::: :::: ::::

```



Db 3172 ACTGTGATGAGCTGCTGACTTACTGTGATTGATTCTAGTCCAGTGGCTTTGTCCTG 3231  
 Qy ArgmetwetiGlyProGluArgGluGlyProProLeuGlyLeuGluLeuLeuAla 1050  
 Db 3232 AAATGATAGGCCCAACCATGAGCCAGATGACAGTACACAAAGCTTGTAAATACGTTAA 3291  
 Qy 1051 GluGlyArgArgLeuProProProProThrcysProThrcGlyValGluGluLeuMetGln 1070  
 Db 3292 GAAGGAAAGCCGCTGCTGCGCCACCTAACTGCATGAGGTATTACAGCTTATGAGA 3351  
 Qy 1071 LeuGlySTPAlaProGluProHisAspArgProAlaPheAlaThrLeu 1086  
 Db 3352 AAATGCTGGGAATTCAACATCGAATGAGACAGCTTTGAGAACCTT 3399  
 RESULT 14  
 AAC66245  
 ID AAC66245 standard; DNA: 3429 BP.  
 AC AAC66245:  
 XX 19-FEB-2001 (first entry)  
 DT 19-FEB-2001 (first entry)  
 DE Human Jak1 kinase nucleotide sequence.  
 XX Jak3; kinase; cytokine; cellular response; inhibition; Jak1; ds;  
 KW cell proliferation; erythrocytosis.  
 XX Homo sapiens.  
 OS  
 XX US6136595-A.  
 PN 24-OCT-2000.  
 PD 18-JUN-1996; 96US-0665574.  
 PF 29-JUL-1994; 94US-0282012.  
 PR 29-JUL-1993; 93US-0097997.  
 PR 09-SEP-1993; 93US-0118968.  
 XX (SUJUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 PA Silvenoinen O, Wiltshuh BA, Ihle J;  
 PI WPI: 2000-666080/67.  
 DR P-PSDB: AAB35725.  
 DR New DNA encoding Jak3 kinase is useful as cytokine regulator for  
 PT treating cell proliferation -  
 PS Disclosure: Fig 2; 100pp; English.  
 XX This invention relates to DNA encoding a murine Jak3 protein. The amino  
 CC acid sequence of the Jak3 protein is given in AAB35725. The Jak family  
 CC of kinases are involved in the cellular response to the binding of  
 CC cytokines to their respective receptors. Jak3 kinase mediated activation  
 CC of some cytokines through their phosphorylation in response to  
 CC cytokine-receptor binding. Inhibiting the activity of Jak3 kinase (at the  
 CC nucleic acid level with antisense sequences or ribozymes, or at the  
 CC protein level with antibodies, kinase inhibitors etc.) is used to treat,  
 CC or diagnose, diseases caused by excessive secretion of certain cytokines,  
 CC e.g. excessive cell proliferation such as erythrocytosis. Alternatively,  
 CC the Jak3 protein and polynucleotide can be used to treat conditions  
 CC associated with defective Jak3 activity. The DNA sequence can be used to  
 CC produce recombinant Jak3 and this used to raise antibodies useful as  
 CC specific inhibitors or to detect or isolate Jak3 without interfering  
 CC with its enzymatic activity. The present sequence represents DNA encoding  
 CC a Jak1 protein used in the isolation and characterisation of the Jak3  
 CC protein of the invention.  
 XX Sequence 3429 BP; 1011 A; 786 C; 856 G; 776 T; 0 other;  
 Alignment Scores: 1.16e-135 Length: 3429  
 Pred. No.:

Score: 1904.00 Matches: 443  
 Percent Similarity: 53.66% Conservative: 188  
 Best Local Similarity: 37.67% Mismatches: 381  
 Query Match: 32.49% Indels: 164  
 DB: 21 Gaps: 24  
 US-09-397-967-16 (1-1099) x AAC66245 (1-3429)  
 Qy 17 SerLeuSerSerSerGluAlaGlyAlaLeuHisValLeu-----LeuProProArgGlyPro 35  
 Db 46 AACCTGGAGGCCCTGAGCCAGAGGGGTGAGAGTATCTTATCTGTGCGAGAGGAGGCC 105  
 Qy 36 GlyProProGluArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuVal 55  
 Db 106 -----CTCCGGCTGGCGCATGAGAGTACAGAGTACAGAGAACTGTGATC 150  
 Qy 56 ArgAlaAlaLysAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThr 75  
 Db 151 AGGGCTGCACAGGATGCGGTATCTCTCTTGTCTCAACACTTTTGGCCCTGTATGAC 210  
 Qy 76 GluAspPheSerCysTyrPheProPheSerHisIlePheCysIleGluAspValAspThr 95  
 Db 211 GAGAACACCAAGCTCTGTGATGCTCCANATGCGACCATCACCCTTGATGACAGATGTC 270  
 Qy 96 GluValLeuValTyrArgLeuArgPheTyrPheProAspTyrPheGlyLeu----- 112  
 Db 271 CTCCGGCTCCACTACCGGATGAGGTATTTTACCATTTGGCATTTGGAAACAGACAAAT 330  
 Qy 113 ---GluThrCysHisArgPheGlyLeuArgLysAspLeu----- 124  
 Db 331 GAGCAGTCAGTGGCGCATTTCTCCAAAGAGAGAAATGGCTACGAGAAAAAAG 390  
 Qy 125 -----ThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141  
 Db 391 ATTCAGATGCAACCCCTCTCTTATGCTCCAGCTACGATGAGATCTTTGCTCAGGGA 450  
 Qy 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln----- 159  
 Db 451 CAGTATGATTGGTGAATGCGCTGCTCTATTCAGACCCCAAGACCCAGACAGAGATGGA 510  
 Qy 160 -----GlyGluPheLeuSerLeuValValLeuAspLeuAlaGlnMetValArg 175  
 Db 511 CATGATTATTGAGAACGAGTGTCTAGGATGCTGCTGCGCATCTCACTATGCCATG 570  
 Qy 176 GluGluAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuPro 195  
 Db 571 ATGAAGAGATGCAAGTTCGCAAGACTGCCCAAGACATCAGCTACAGCATATATTCGA 630  
 Qy 196 ProSerLeuArgAspValIleGlnGlyGlnAspPheValThrArgArgGlyLeu----- 213  
 Db 631 GAACATTTGAATTAAGTCCATCAGACAGAGAACTTTCACCCAGGATCGGATTAATAT 690  
 Qy 214 -----ArgArgThrValValLeuAlaLeuLeuPro 223  
 Db 691 GTTTTCAAGCATTTCTTAAGAGAAATTTAAACAACAGACATTTGTGACAGCAGCTG 747  
 Qy 224 CysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGlu 243  
 Db 748 -----TCCACGATGACCTGAGAGGATGAAATATCTTGCTACCTTGGA 789  
 Qy 244 ArgLeuHisProAlaAlaIleThrIleGluThrPheArgValGlyLeuProGlyAlaGlnGlu 263  
 Db 790 ACTTTCACAAACATTTAGCGTGTCTGTAATATTT----- 822  
 Qy 264 GluProGlyLeuLeuArgValAlaGlyLysAspGlnGlyIleProTyr---SerSerAsp 282  
 Db 823 GAGACTTCAGTGTACTGATTTTCATCAGAAATGAGATGAATGTGTTTCATTCGAATAC 882  
 Qy 282 ----- 282  
 Db 883 GGTGAACGTTCTCTACTAGAGATGATGTTGACTGGGAATCTTGGAATCCAGTGGAG 942  
 Qy 282 ----- 282

Db 943 CATAAACCAATGTGTCTTCTGAAAGAAAAAATAACTGAAGCGGAAAAAATCTG 1002  
QY 283 -----GluLeuPheGlnThrPhe 288  
Db 1003 GAAATATAAGACAGAAGATGAGAGAAAAACAGATCCGGAGATGAGCAATTTT 1062  
QY 289 CysAspPheProGluLeuValAspValSerLLeuSnglnAlaProArgValGlyProAla 308  
Db 1063 TCATTCTCTCCCTGAATCACTCATCTGTAATTAAGAGACT----- 1104  
QY 309 GlyLlGlnHisArgLeuValThrValThrArgMetAspGlyHisLLeuGlnAlaGluPhe 328  
Db 1105 -----GTGGTCACACATTAAACACAGCAGACACAAAGAAATGAGAACTGAACCTC 1152  
QY 329 ProGlyLeuProGlnAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIle 348  
Db 1153 TCTTCCACAGAGAGCGCTGCTTGTCTTGTCTCCCTGTAGATGGCTACTCCGGCTCACA 1212  
QY 349 CysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuGlnGluGlu 368  
Db 1213 GCAGATGCGCATATTACTCTGTGACCGAGCTGCCCCGTTGATGCTGCACACACATA 1272  
QY 369 AlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAla 388  
Db 1273 CAGAAATGGCTGCATGGTCCAACTCTGTACAGAAATAGCCCATCAATTAATTGGCGCAAGAA 1332  
QY 389 GlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeu 408  
Db 1333 GGAAGCGAGAGGAGGATGATGCTGAGTGAGTGAGCTGCACCACTTTTGACACACTCTC 1392  
QY 409 LeuThr-----AlaCysValGlnThrPro-----LeuGlyProAspTyrLys 422  
Db 1393 ATGACCGCTCACCTGCTTGTAGAGTGTGACAGAGTGTGACGGGTGCGCGAAGCACTTAAG 1452  
QY 423 GlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnPro 442  
Db 1453 AACTTCAGATC---GAGGTGCGAAGAGCGCGTACAGTGTGACAGGTTGCGACCGCAGC 1509  
QY 443 HisArgSerLeuArgGluLeuLeuAlaAlaCysTyrAsnSerGlyLeuArgValAspGly 462  
Db 1510 TTCCCCAGCGTTGGAGAGCTCATGAGCCACCTCAGACAGACAGATCTGCGCACGGATAAC 1569  
QY 463 AlaAlaLeuTyrLeuThrSerCysAlaProArgProLysGlnLysSerAsnLeuIle 482  
Db 1570 ATGAGCTTCAATGCTAAAGCTGTGCTGCCACCAAGCCCGAGAAATCTCCAACTCTCTG 1629  
QY 483 ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAla 502  
Db 1630 GTGGCTACTTAAG-----AAAGCCAGAGATGGCAGCCCC---GTCTACCCC 1671  
QY 503 LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGlnTyrPHisGlnAsnLeu 522  
Db 1672 ATGAGCCAGCTGAGTTTCATCGGATCTCAGAAAGATGTGGTGCAGCGGCGAGCACTT 1731  
QY 523 GlyHisGlySerPheThrLysIlePheArgGlySerArgArgLysValValAspGlyGlu 542  
Db 1732 GGGAGAGGACGAGAGAACACATCTATTCTGGAGCCCTGATGATGTCAAGATGAGGAA 1791  
QY 543 ThrHisAspSerGlu-----ValLeuLeuLysValMetAspSerArgHisArg 558  
Db 1792 GGAACCTTCTGAAGAGAAAGATAAAGTATCTCAAGCTTAAAGACCCAGCCAGCAGG 1851  
QY 559 AsnCysMetGluSerPheLeuGlnAlaAlaSerLeuMetSerGlnLysSerTyrProHis 578  
Db 1852 GATATTTCCCTGGCTTCTTCGAGCGAGCCAGCATGTGAGACAGGCTCTCCACAAACAC 1911  
QY 579 LeuValLeuLeuHisGlyValCysMetAlaGly---AspSerIleMetValGlnGluPhe 597  
Db 1912 ATCGTGTACTCTATGCGCTGTGTGTCGAGCGTGTGAGAAATATCATGTGTGAAGACTTT 1971  
QY 598 ValTyrLeuGlnAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSer 617  
Db 1971 -----

Db 1972 GTGGAAGGGGCTCCTCGATCTCTTCATGACCGGAAAAAGTATGTCCTTACACACCA 2031  
QY 618 TrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspLysGly 637  
Db 2032 TGGAAATTCAAAGTTGGCAACAGCTGCGCAGTCCCTGTGACTGAGTGGAGATTAAGAC 2091  
QY 638 LeuProHisGlyValSerAlaArgLysLeuLeuAlaArgGlnGluGlyAspGly 657  
Db 2092 CTGGTCCATGGAATGTGTACTATAAAACCTCTCTCTGCGCCGTGAGGAAATCGACAGT 2151  
QY 658 Asn---ProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu 676  
Db 2152 GAGTGTGGCCATTCATCAAGCTCAGTACGCCCGCATCCCATTCAGGTGCTGTAGG 2211  
QY 677 GluMetLeuThrAspArgIleProTyrPValAlaProGlyCysLeuGlnAlaGlnThr 696  
Db 2212 CAAGAAATGATTAAGAACCAATCCATGATGCTCCCTGAGTGTGAGGACTCAAAAC 2271  
QY 697 LeuCysLeuGlnAlaAspLysTyrPValPheGlyValaThrThrTrpGluValPheGlnArg 716  
Db 2272 CTGAGTGTGGCTGTGACAGTGTGAGCTTGTGAAACACGCTGTGGAAATCTGCTACAT 2331  
QY 717 GlyProAlaHisIleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGlnAspGln 736  
Db 2332 GCGGAGATCCCTTGAAGAGCAAGACGCTGATTGAGAAAGAGATTCATGAAGCGCG 2391  
QY 737 GlyGlnLeuProAlaLeuLysTrpThrGlnLeuAlaGlyLeuIleThrGlnCysMetAla 756  
Db 2392 TGCAGCGCATGACACCATATGTAAGAGACTGTGACTACCTATGACCCCTGATGAC 2451  
QY 757 TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle 776  
Db 2452 TATGACCCCAATCAGAGAGCGCTTCTTCGAGCATCATGAGACATATTAAGCTTGAA 2511  
QY 777 ThrSerAspTyrGluLeuLeuSer-----AspProThrProGlyIleProSer 792  
Db 2512 GAGCAGAGATCAGATATGTTTCCAGAAAAAACCAGCCACTGAA----- 2559  
QY 793 ProArgAspLysLeuValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIle 812  
Db 2560 -----GTGACCCCAACCAT 2574  
QY 813 PheGlnGluArgHisLeuLysTyrIleSerLeuGlnGlyLysGlnAsnPheGlySerVal 832  
Db 2575 TTGAGAGCGCTTCTTCAAGAGAGATCCGTGACTTGGAGAGGCGCTTGGGAAGGTT 2634  
QY 833 GluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGln 852  
Db 2635 GAGCTGTGAGGTATGACCCC---GAAGACAAATACAGGAGCAGGAGGCTGTAAATCT 2691  
QY 853 LeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgLysIleGlnLeu 871  
Db 2692 CTGAGCGCTGAGAGTGAAGGTAAACCATAGCTAGCTGAAAGAAATCGAGATCTTA 2751  
QY 872 LysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArg 891  
Db 2752 AGGAACTCTATCATGAGCACTGTGAATACAAAGAAATCTGCACAGAAAGCGGAGA 2811  
QY 892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyLysLeuArgAspLeuGln 911  
Db 2812 AATGCTATTAACTCATCATGGAATTTCTGCTTGGGAAGCCCTTAAGAAATCTTCCA 2871  
QY 912 ArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLys 930  
Db 2872 AAGATATAAGAAAAAATAAATCAACCTCAACAGACAGATAAATATGCGCTTGAATTAAG 2931  
QY 931 GlyMetGluTyrLeuGlnGlyAlaArgArgCysValHisArgAspLeuAlaIleArgAsnIle 950  
Db 2932 GGGATGACATATTGGTGTCTGCGCAATACCTTCACCGGAGACTTGGCAGAGAAATGTC 2991  
QY 951 LeuValGlnSerGlnAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuPro 970  
Db 2992 CTGTGTGAGGTGAAACCAAGTGAATAATGAGACTTCCGTTTAACCAAGCAATTTGAA 3051







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 28, 2003, 16:07:34 ; Search time 302.323 Seconds

(without alignments)  
3955.543 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 5860  
Sequence: 1 MAPSEPTPLIPORCSLSS.....RPAFATLSPQLDPLMRGRP 1099

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 709820 seqs, 544064369 residues

Total number of hits satisfying chosen parameters: 1419640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool/US09397967/Runat\_26042003\_182314\_10830/app.query.fasta\_1.1486  
-DB=Published\_Applications\_NA -QW=fastap -SUFFIX=rpnb -MINMATCH=0.1  
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-NCPU=6 -ICPU=3 -NO\_XLERY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT  
-DSPBLOCK=100 -LONLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THRADS=1 -XGAPOP=10  
-XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications\_NA:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCIT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCITUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1908.5	32.6	4191	9	US-09-962-854A-1
2	1904	32.5	3541	9	US-09-880-107-2379
3	1842	31.4	3538	9	US-09-962-854A-2
4	713	12.2	2338	10	US-09-925-302-90

5	609.5	10.4	2027	10	US-09-771-161A-29	Sequence 29, Appl
6	483.5	8.3	975	9	US-09-962-854A-5	Sequence 5, Appl1
7	482.5	8.2	753	10	US-09-910-943-636	Sequence 636, App
8	482	8.2	3914	12	US-10-044-090-148	Sequence 148, App
9	475	8.1	3768	10	US-09-811-123-8	Sequence 8, Appl1
10	475	8.1	3768	10	US-09-811-115-2	Sequence 11, Appl1
11	475	8.1	4530	9	US-09-877-177-11	Sequence 11, Appl1
12	475	8.1	9274	10	US-09-811-123-7	Sequence 7, Appl1
13	475	8.1	3768	9	US-09-854-356-9	Sequence 9, Appl1
14	474	8.1	3768	9	US-09-930-125-1	Sequence 1, Appl1
15	474	8.1	4473	9	US-09-441-411-5	Sequence 1, Appl1
16	474	8.1	4543	9	US-09-763-508-1	Sequence 1, Appl1
17	471	8.0	3633	10	US-09-725-433-1	Sequence 1, Appl1
18	466.5	8.0	5484	10	US-09-940-101-1	Sequence 98, Appl1
19	465.5	7.9	10058	9	US-09-974-298-98	Sequence 10, Appl1
20	464.5	7.9	3955	9	US-09-854-356-10	Sequence 10, Appl1
21	464.5	7.9	3955	9	US-09-870-759-117	Sequence 11, Appl1
22	464.5	7.9	3955	9	US-09-854-356-11	Sequence 11, Appl1
23	458.5	7.8	3771	9	US-09-854-356-11	Sequence 19, Appl1
24	455	7.8	7607	10	US-09-962-610-19	Sequence 1731, Ap
25	450.5	7.7	3726	10	US-09-925-302-271	Sequence 271, App
26	448.5	7.7	2437	9	US-09-158-722-3	Sequence 48, Appl1
27	448.5	7.7	3454	12	US-10-044-090-48	Sequence 1631, Ap
28	442.5	7.6	2863	10	US-09-954-456-1631	Sequence 6, Appl1
29	441	7.5	1755	9	US-09-930-125-6	Sequence 4, Appl1
30	441	7.5	1767	9	US-09-930-125-4	Sequence 7, Appl1
31	441	7.5	1773	9	US-09-930-125-7	Sequence 5, Appl1
32	441	7.5	1806	9	US-09-930-125-5	Sequence 1731, Ap
33	438	7.5	5264	10	US-09-920-300A-1731	Sequence 5, Appl1
34	438	7.5	5264	12	US-10-033-528-1731	Sequence 1731, Ap
35	437.5	7.5	2770	9	US-09-977-260-5	Sequence 5, Appl1
36	437.5	7.5	2770	9	US-09-977-261-5	Sequence 5, Appl1
37	437.5	7.5	2770	10	US-09-977-269-5	Sequence 5, Appl1
38	428.5	7.3	4879	9	US-10-172-620-13	Sequence 13, Appl1
39	428.5	7.3	4975	10	US-09-880-107-2342	Sequence 2342, Ap
40	425	7.3	2454	10	US-09-964-824A-120	Sequence 120, App
41	425	7.3	2627	10	US-09-962-436-313	Sequence 313, App
42	420.5	7.2	4541	9	US-10-123-036-3	Sequence 3, Appl1
43	420.5	7.2	4541	10	US-09-880-107-3785	Sequence 3785, Ap
44	419.5	7.2	3060	10	US-09-824-859A-6	Sequence 6, Appl1
45	418	7.1	2674	9	US-10-003-295-1	Sequence 1, Appl1

## ALIGNMENTS

RESULT 1  
US-09-962-854A-1  
Sequence 1, Application US/09962854A  
Publication No. US20030044803A1  
GENERAL INFORMATION:  
APPLICANT: Pedersen, Finn S.  
APPLICANT: Hernandez, Javier Martin  
TITLE OR INVENTION: Methods for diagnosis and treatment of diseases associated with  
FILE REFERENCE: A-70020/RMS/DCF  
CURRENT APPLICATION NUMBER: US/09/962, 854A  
PRIOR FILING DATE: 2001-09-24  
PRIOR APPLICATION NUMBER: US 09/668, 644  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 4191  
TYPE: DNA  
ORGANISM: Mus sp.  
US-09-962-854A-1

## Alignment Scores:

Pred. No.: 2,29e-191  
Score: 1908.50  
Percent Similarity: 54.52%  
Best Local Similarity: 38.37%  
Length: 4191  
Matches: 437  
Conservative: 184  
Mismatches: 383

Query Match: 32.57% Indels: 135  
 DB: 9 Gaps: 22

US-09-397-967-16 (1-1099) x US-09-967-854A-1 (1-4191)

41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValAlaGlyAlaIleValAla 60  
 408 CTCGGCTGGGACGAGAGTATACAGCCGAGAGCTGTCATCGAGGCCGCCGAGAG 467  
 61 CysGlyIleLeuProValTyrHisSerPheAlaLeuAlaThrGluAspPheSerCys 80  
 468 TGCAGTATCTCTCTCTCTCAACCTCTTGCCCTGACATGAGACACCAGCTC 527  
 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGluValLeuValTyr 100  
 528 TGGTACGCTCGGACGAGCATCATCATCATGATGACAAAGCTCTCCGCTCCACTAC 587  
 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeu-----GluThrCysHis 116  
 588 CGCATGAGGTTCTACTTACCACTGGCAGGAAACAATGACAAACAGACAGTCTGTGG 647  
 117 ArgPheGlyLeuArgLysAspLeu-----ThrSer 126  
 648 GCATCTCTCCAAAGAACGAAAGAAAGCGCTATGAGAAAGAGTTCCAGAACCAAC 707  
 127 AlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHisArgSerAspLeuVal 146  
 708 CCATCTCTTGTATGCCAGTTCATGAGATCTGTTTGCACAGGAGCAGATGATTTGATC 767  
 147 SerGlyArgLeuProValGlyLeuSerMetLysGluGln-----Gly 160  
 768 AAATGCCGCTCCCATTCGGGACCCCAAGACGAGGACGACATGATATTGAAAT 827  
 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
 828 GAGTCCCTGGGACATGGCGCTCTGCGCATTCGCCATATGCCATGATGAAACAGATCGAG 887  
 181 ProGlyIleLeuLeuLysThrValSerTyrLysLysLysLeuProPheSerLeuArgAsp 200  
 888 TTGCGGAGACTCCCAAGACATCGACTACAGCATATATCCAGAAACATTTGAATAA 947  
 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIle----- 213  
 948 TCCATTCAGACAGAGAACCTTCTTACAGAGATGCGAATAAATGTTTCAAGANTTC 1007  
 214 -----ArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeuPro 228  
 1008 TTGAAGGAATTTAACACAGACCATC-----TGTGAC 1040  
 229 GlyArgProTyrAlaIleMetAlaLysTyrIleLeuAspLeuGlu-----ArgLeuHis 246  
 1041 AGCATGTGTCATGACCTGAGAGTGAATATACCTGCTACCTTGGAAATCTTACATTGACA 1100  
 247 ProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGluGluProGly 266  
 1101 AAACATTTATGAGCTGAATATATTGAGACTTCTATGCTACTGATTTTCATCGAAATGAA 1160  
 267 LeuLeuArg-----ValAlaGly 272  
 1161 TTGAGTTCGATGCCATTCGATGACAGTGCATGTTCTCTATGAGTCAATGAGTACTGGA 1220  
 273 AspAsnGlyIleProTyrPheSerSerAsnAsp----- 282  
 1221 AATTCCTGGATCCAGTGGCGGAGAAACCAATGTTGTTCTGTTGAAAGGAAAAAAT 1280  
 282 ----- 282  
 1281 AAATGACGCGGAAAAAAGCTGCAATTAATAACACACAGAGATGACAGAAACAAA 1340  
 283 -----GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsn 300  
 1341 CTCGGGGAAGAGTGAACAATTTTCTTATTTCCCTGAATATCACCCACATCTGTAATAAAG 1400

301 GlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAsp 320  
 1401 GAGTCT-----GTGTCACGATTAACAAACAGAGC 1430  
 321 GlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeu 340  
 1431 AACAAACATGAGACTCAAGCTCTCTCGAGAGAGAGCTTCTCTTGTGTCTCTG 1490  
 341 ValAspGlyTyrPheArgLeuIleCysAspSerAlaGlySerPheCysLysGluValAla 360  
 1491 GTGATGCTTACTTCGGCTCACTGACAGATGCCACCATATACCTGTGTACTGATGTGCT 1550  
 361 ProProArgLeuLeuGluGluGluAlaAspValCysHisGlyProIleThrLeuAspPhe 380  
 1551 CCCCCCATATTTGTCACAAATATACAGAAAGCGCTCCAGCGTCAATGTCACAGAAAT 1610  
 381 AlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSer 400  
 1611 GCATCAATATACCTGCGCGAGAGAGGAGTGAAGAGGATGTACTGTGAGTGGAGC 1670  
 401 ProGluAspTyrAspSerPheLeuLeuThr-----AlaCysValGlnThrPro-----Leu 417  
 1671 TGCACCGACTTGTACACATCTTATGACTGTACCTGCTTGGAAAGTGTGAGTATGG 1730  
 418 Gly-----ProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSer 435  
 1731 GTGGCCAGACAGCATTCACAAACTTTCAGATT--GAGGTACAGAAAGGCCGTACAGC 1787  
 436 LeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuAlaIleValCysTrpAsn 455  
 1788 CTCATGCTCTATGACACCATTTCCACCTGCGAGACTCATGATCAACCACTCAAGAG 1847  
 456 SerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgPro 475  
 1848 CAGATCCTGGCGACGAGCAACATATACCTTGTCTGTAACGATGCTGTAGCTTAAGCT 1907  
 476 LysGluLysSerAsnLeuIleValAlaArgArgGlyCysAsnProAlaProAlaProGly 495  
 1908 CGAAGATTCCTCAATCTGCTGAGCCACTAAG-----AAAGCCAGAGG 1952  
 496 CysSerProSerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSer 515  
 1953 TGGCAGCT---GTCATCTCATGAGCAGCTGAGCTTGATGATGCTTAAAGAAAGAT 2009  
 516 LeuGluThrPheIleAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArg 535  
 2010 ATTATACAGGTGAGCACCTTGGCAGAGCAAGAACATATCTATCTGTGGACCTG 2069  
 536 ArgGluValValAsp-----GlyGluThrHisAspSerGluValLeuLeuLys 551  
 2070 CTGGACTACAGATGAGAGCAATTCGTGAAGAGAAAGATTAAGATGATCTCAAA 2129  
 552 ValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaIleAspLeuMet 571  
 2130 GTCTTACAGCCACCGACCGGACATCTCTGCGCTCTTCTTAAAGGTGTGTACTACATGATG 2189  
 572 SerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGly---Asp 590  
 2190 AGACAGGTTTCCCAACAAATATAGTGTACTACAGCGGTGTGTCCAGATGTGGAA 2249  
 591 SerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArg 610  
 2250 AATATCATGTGGAAGATTTGTGAGGGGGCGCTTGTGATCTCTTCAATGACCGGAAA 2309  
 611 GlyHisLeuValSerAlaSerTyrLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeu 630  
 2310 AGTATGCGCTTACTACCCCTGGAAGTCAAGTGTCCAAACAGCTGCCAGAGCGCTG 2369  
 631 AsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeu 650  
 2370 ACTTACTTGGAAAGATTAAGACCTGTTCTATGTAATGTGTCACATAAAGCTCTCTCTG 2429  
 651 AlaArgGluGlyGlyAspGlyAsn---ProProHelleuLysLeuSerAspProGlyVal 669



```

Db 2420 GCGCGTAGGCGCATGACATGATGCGCGCTTCATCAAGCTTAGAGCCCTGCGATC 2489
Qy SerProthrinLeuSerLeuGluMetLeuThrAspArgIleProTyrValAlaProLu 689
Db 2490 CCACTCTGTGGTGCAGGAGGAGGAGTGCATAGAGCAATCCCTGATGCTCTTAG 2549
Qy CysLeuGlnGluAlaGlnThrLeuGluGluAlaAspLysTyrPglPheGlyValAlaThr 709
Db 2550 TGTGTGAAGACTCCAAACACCTGAGCTGTGCTGCAAGTGGAGCTTTGGAAACGCG 2609
Qy ThrTyrPglValAlaPheGluArgGlyProAlaHisIleThrSerLeuGluProAlaLysLys 729
Db 2610 CTCTGGGAAATCTCTACACAGAGAGATTCCTCAAGAACAGACCCCTATTGGAAA 2669
Qy LeuLysPheTyrGluAspGlnGlyGluLeuProAlaLeuLysTyrThrGluLeuAlaGly 749
Db 2670 GAGAGGTTTATGAAAGCCGCTGCAGGCTGTGACTCCATCTTGCACAGGAGCTGAC 2729
Qy LeuIleThrGlnCysMetAlaTyrAspProGluArgArgProSerPheArgAlaIleLeu 769
Db 2730 CTCATGATCGCTGCTGATGACTATGACCCCAACAGAGACCCCTCTTGCAGCCATCATG 2789
Qy ArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspProThrProGly 789
Db 2790 AGGACATTAACAAGCTGAGAGAGAGAGATTCAGACATGTTTCAGAA----- 2837
Qy IleProSerProArgAspGluLeuGluCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAsp 809
Db 2838 ---AAGCAGCCACACAGAGAGTGC-----GAC 2861
Qy ProAlaIlePheGlnGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPhe 829
Db 2862 CCCACTCACTTTGAAAGCGGCTTCCTGAAGAGATTCGTGACTGGAGAGGCTCAGCTT 2921
Qy GlnSerValGluLeuGluCysArgTyrAspProLeuGlnAspAsnThrGlyProLeuValAla 849
Db 2922 GGGAGAGTTGAGCTCTGAGATATGATCTCAGAGAGACACAGAGGAGCAGTACT 2981
Qy ValLysGlnLeuGln---HisSerValProAspGlnGlnArgAspPheGluArgGluIle 868
Db 2982 GTCAAGTCCCTGAGAGCCGTGAGAGTGGAGGTAACCATAGCTGATCTGAAGAGAGATA 3041
Qy GlnIleLeuValAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGly 888
Db 3042 GAGATCTTACGGAACCTCTACCATGACAAATTTGGAAGTCAAGAAATTCGATGAA 3101
Qy ProGluArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAsp 908
Db 3102 GACGAGGCAATGCTATGACCTCTCATGAGATTCTGCTTGGGAGAGCCCTAAAGAG 3161
Qy LeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTyrPgl 927
Db 3162 TATCTGCCAAGATAGAAACAAATCAACCTCAACAGCAGCAATTAATATGATCCATCAG 3221
Qy IleCysLysGlyMetGluTyrLeuGlnValAlaArgCysValHisArgAspLeuAlaAla 947
Db 3222 ATTTGTAAGGGAGTGAATGCTTGTGCTTGTGCAATACCTGCGGAGCTTATGACACA 3281
Qy ArgAsnIleLeuValGlnSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLys 967
Db 3282 AGAAATGCTCTGTGTGAGAGGAGCATCACTGAAGATCGGAGACTTGTGTTAAACCAA 3341
Qy LeuLeuProLeuGlnLysAspTyrTyrValValArgGluProGlnGlnSerProIlePhe 987
Db 3342 GCATTTGAACCGATAGAGACTACTACACATCAAGAGCAGCAGGAGACGCCAGCTGTC 3401
Qy TrpTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrPser 1007
Db 3402 TGGTACGCTCCGGAATGTTTATTCAGTGTAAATTTATATGCGCTGATGCTGCTGCT 3461
Qy PheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027

```

```

Db 3462 TTGGAGTGCATGCTGCAGAGCTGCTACTTACTGTGACTGAGATTTTAGTGCATGCC 3521
Qy GluPheLeuAlaGlyMetGlyProGluArgGluGlyProProLeuGluCysArgLeuLeuGlu 1047
Db 3522 TTGTCTCGAAATAGATGAGGCGCAATCATGAGCCAGATGACAGTGCAGCGCTGTGAG 3581
Qy LeuLeuAlaGluGlnArgArgLeuProProProProThrGlnValGlnGlu 1067
Db 3582 ACTCGAAGAGAGCAACGCTGCTGATGCTCACCACCACTGCTCTGATGAGAGTTATGAG 3641
Qy LeuMetGlnLeuGlyTyrAlaProGluProHisAspArgProAlaPheAlaThrLeu 1086
Db 3642 CTTATGAGAAATGCTGGAATTCACACATCTAACCGAGACACTTTCAGAACCTT 3698

RESULT 2
US-09-880-107-2379
: Sequence 2379, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-NO
: CURRENT APPLICATION NUMBER: US/09/880,107
: PRIOR FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn ver. 2.1
: SEQ ID NO 2379
: LENGTH: 3541
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 M64174
US-09-880-107-2379

Alignment Scores:
Pred. No.: 5,32e-191 Length: 3541
Score: 1904.00 Matches: 443
Percent Similarity: 53.66% Conservative: 188
Best Local Similarity: 37.67% Mismatches: 381
Query Match: 32.49% Indels: 164
DB: 10 Gaps: 24

US-09-397-967-16 (1-1099) x US-09-880-107-2379 (1-3541)
Qy 17 SerLeuSerSerGluAlaGlyAlaLeuHisValLeu---LeuProProArgGlyPro 35
Db 121 AACCTGAGAGGCGCCCTGAGCCAGGAGTGAATGATCTTCTTGTCTGCGACAGGAGGCC 180
Qy 36 GlyProProGlnArgLeuSerPheSerPheGlnAspTyrLeuAlaGluAspLeuGlyVal 55
Db 181 -----CTCGGCTGGGAGCGAGTACGAGATGACACAGCAGAGGAAGAACTGTCATC 225
Qy 56 ArgAlaAlaLysAlaGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThr 75
Db 226 AGGCTGACAGGAGATGCGCTATCTCTCTTGTGTCAACACTTGTGCGCTGTATGAC 285
Qy 76 GluAspPheSerCysTyrPheProProSerHisIlePheCysIleGluAspValAlaPthr 95
Db 286 GAGAACACCAAGCTCTGTGTATGCTCCAAATCGCACCATCACCGTTGATGAGACAATGTC 345
Qy 96 GlnValLeuValTyrArgLeuArgPheTyrPheProAspTyrPheGlyLeu----- 112
Db 346 CTCGGGCTCCACTACCGGATGAGGTCTATTTCAACATGTCATGAGAACAGACAGCAAT 405
Qy 113 ---GluThrCysHisAspPheGlyLeuArgLysAspLeu----- 124

```

QY	409	LeuThr---AlaCysValGlnThrPro-----LeuGlyProAspTyrLys	422
		:::     :::     :::     :::	
Db	1468	ATGACCCGTCACCTGGCTTTGGAAGTCTGACAGGTGACAGGTCAGGTCGCCAAGACAGTTCACAG	1527
QY	423	GlyCysLeuLeuArgGlnAspProSerGlyAlaIlePheSerLeuValGlyLeuSerGlnPro	442
		:::     :::     :::     :::     :::     :::     :::	
Db	1538	AACTTTCAGATC---GAGGTGCGAAGAGGCCCGCTACAGTCTCCACCGTTTCGGACCCGAC	1584
QY	443	HisArgSerLeuArgGlnLeuLeuAlaAlaCysTyrPasnSerGlyLeuArgValAspGly	462
		:::     :::     :::     :::     :::     :::     :::	
Db	1585	TTCCCCACGCTTGGAGACCTCATGACCCACCTCAAGAACATCATCTCCGACGAGTAAAC	1644
QY	463	AlaAlaLeuTyrLeuThrSerCysAlaProArgProGlyGlnLysSerAsnLeuIle	482
		:::     :::     :::     :::     :::     :::     :::	
Db	1645	ATCGAGCTTCATGCTAAACCGCTCTCCAGCCCAAGCCCGCGAAGATCTCCACCTGGTG	1704
QY	483	ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAsla	502
		:::     :::     :::     :::     :::     :::     :::	
Db	1705	CTGGCTACTAAG-----AAAGCCAGAGATGGACGCC---GTCCACCC	1746
QY	503	LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGlnThrHisGlnAsnLeu	522
		:::     :::     :::     :::     :::     :::     :::	
Db	1747	ATGACCCAGCTGAGTTCGATCGGATCTCCAAAGAGATCTGTCACAGGCGACACTT	1806
QY	523	GlyHisGlySerPheThrLysIlePheArgGlySerArgArgGlnValValAspGlyGln	542
		:::     :::     :::     :::     :::     :::     :::	
Db	1807	GGGAGAGCGACAGAACACATCATCTTCGGGACCTGATGATTACAGAGATGACGAA	1866
QY	543	ThrHisAspSerGlu-----ValLeuLeuLysValMetAspSerArgHisArg	558
		:::     :::     :::     :::     :::     :::     :::	
Db	1867	GGAACCTCTCGAAGGAGAGATAAAGATGATCTTCAAGTGTGAGCCCGCCACAGG	1926
QY	559	AsnCysMetGlnSerPheLeuGlnAlaIleSerLeuMetSerGlnValSerTyrProHis	578
		:::     :::     :::     :::     :::     :::     :::	
Db	1927	GATATTTTCCCTGGCCTTCTTCGAGGACGACGAGATGATGACAGAGTCTCCCAACAC	1986
QY	579	LeuValLeuLeuHisGlyValLysMetAlaGly---AspSerIleMetValGlnIlePhe	597
		:::     :::     :::     :::     :::     :::     :::	
Db	1987	ATCGGTACCTCTATGGCGTCTGTGCCGACGCTGGGAAGAAATCATGCTGGGAAGATT	2046
QY	598	ValIleuLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSer	617
		:::     :::     :::     :::     :::     :::     :::	
Db	2047	GTGCAAGGGGCTCTCTGCGATCTCTTCATGCAACGGAAAGATGCTCTTACCAACACA	2106
QY	618	TrrPylsLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspLysGly	637
		:::     :::     :::     :::     :::     :::     :::	
Db	2107	TGCAAAATTCAAAGTTGCCCAACAGCTGGCCAGTCCCTGAGCTACTTGAGAGATTAAGAC	2166
QY	638	LeuProHisGlyAsnValSerAlaIleArgLysValLeuLeuAlaArgGlnGlyValAspGly	657
		:::     :::     :::     :::     :::     :::     :::	
Db	2167	CTGGTCCATGGAAATGTGTCTACTAAACACTCTCTGCGCCGTGAGGGAATGCACAGT	2226
QY	658	Asn---ProPheIleLeuLysLeuSerAspProGlyValSerProThrValLeuSerLeu	676
		:::     :::     :::     :::     :::     :::     :::	
Db	2227	GAGTGTGGCCATTCATTCACAGCTCAGAGACCCGCGATCCCATTAACGGTGTGCTCAG	2286
QY	677	GlnMetLeuThrAspArgIleProTrrPylAlaProGlnLysLeuGlnGlnAlaGlnThr	696
		:::     :::     :::     :::     :::     :::     :::	
Db	2287	CAAGATTCATTCAGCAAAATCCATGATGATGCTCCGATGTTGTGAGACATCCCAAGAC	2346
QY	697	LeuCysLeuGlnAlaAspLysTrrPylGlyPheGlyAlaThrThrTrrPylValPheGlnArg	716
		:::     :::     :::     :::     :::     :::     :::	
Db	2347	CTGAGTGTGGCTGTGACAGTGAAGTGGATTTGGAAACACGCTGTGGAAATGTGCTACAT	2406
QY	717	GlyProAlaHisIleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGlnAspGln	736
		:::     :::     :::     :::     :::     :::     :::	
Db	2407	GCGCGAGATCCCTTGAAAGACAGACGCTGATTGAGAAAGAAAGATTTATGAAAGCCGG	2466
QY	737	GlyGlnLeuProAlaLeuLysTrrPylGlnLeuAlaGlyLeuIleThrGlnCysMetAla	756
		:::     :::     :::     :::     :::     :::     :::	
Db	2467	TGCAGGCCAGTGACACCATATGTAAGAGACTGTGGCTGACCTATAGACCGCTGCATGAC	2526



```
Db 765 GTTTCAGAGATTTCCTTAAGAAATTACACAGACCATTTCGTGACGACGCTG--- 821
Oy 224 CysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaTyrTyrIleLeuAspLeu 243
Db 822 -----TTCACGCATGACCTGAAGGTGAAGTAATCTGTGCTACCTGGAA 863
Oy 244 ArgLeuHisProAlaIleThrThrGluThrPheArgValGlyLeuProGlyAlaGln 263
Db 864 ACTTTGACAAAACATTACGCTGCTGAATATTT----- 896
Oy 264 GluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyr---SerSerAsn 282
Db 897 GAGACTTCATGTTACGTATTCATTCACAAATGATGATGATGTTGTTTCATTCGATAC 956
Oy 282 ----- 282
Db 957 GGTGGAACGTTCTCTACTACGAAAGTGTGACTGGAAATCTGCAATCCAGTGAG 1016
Oy 282 ----- 282
Db 1017 CATAAACCAATGTGTTCTGTTGAAAAAGAAAAATAACTGAAAGCGGAAAAATCG 1076
Oy 283 -----GluLeuPheGlnThrPhe 288
Db 1077 GAAATATAACACAGAGATGAGAGAAAAACAAGATCCGGAAAGAGTGGAACAATTT 1136
Oy 289 CysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro 308
Db 1137 TCTTACTCCCTGAAATACACACATTTGTAATAAGAGCT----- 1178
Oy 309 GlyIleHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu 328
Db 1179 -----GtGtGACGATTACACAGACAGACACAAAGAAATGGAATGAGCTG 1226
Oy 329 ProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 348
Db 1227 TCTTCCACGAGAGGAGGCTTGTCTTGTCTGCTGCTGAGTACGCTACTTCGCGCTCA 1286
Oy 349 CysAspSerArgHisTyrPheCysGlyValAlaProProArgLeuLeuGluGlu 368
Db 1287 GCAGATGCCCATCTACCTCTGCACGACGCGGCCCGCTGATGCTGCACACACTA 1346
Oy 369 AlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisIleLeuAla 388
Db 1347 CAGATGCTGTCATGTCATCATCTGTCACAGATACGCCATCAATTAATTTGGGCAAGA 1406
Oy 389 GlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPhe 408
Db 1407 GGAAGCGAGGAGGATCTAGCTGTGAGTGGGCTGC-ACCAGCTTTGACAAACATCTC 1465
Oy 409 LeuThr---AlaCysValGlnThrPro-----LeuGlyProAspTyrIle 422
Db 1466 ATGACCGTCACCTGTTTGAGAGAGTCTGACGACGTCAGGCGGCCACAGACAGTTCA 1525
Oy 423 GlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGln 442
Db 1526 AACTTCAGATC---GAGTGTGAGAGGCGGCTACAGTCTGACAGGTTGCGACGCGAG 1582
Oy 443 HisArgSerLeuArgGlnLeuLeuAlaIleCysTrpAsnSerGlyLeuArgValAsp 462
Db 1583 TTCCCACTGCTGGAGACCTCATGAGCCACTCAAGAAAGACAGATCTCGCAGGATAC 1642
Oy 463 AlaAlaLeuTyrLeuThrSerCysAlaProArgProGlyLeuSerAsnLeuIle 482
Db 1643 ATCAGCTTCATGTAAGACGCTGCTGCCAGCCAGCCCGGAAATCTCCACACTGCTG 1702
Oy 483 ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAla 502
Db 1703 GTGGCTACTAAG-----AAAGCCAGAGTGGCAGCCC---GTCATCCCC 1744
Oy 503 LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluThrHisIleAsn 522
Db ----- 522
Db 1745 ATGAGCCAGCTGATTTGCATCGATCTCTCAAGAAGAGATCTGTTGACGCGACACCT 1804
Oy 523 GlyHisGlySerPheThrLysIlePheArgGlySerArgArgValAlaAspGly 542
Db 1805 GGGAGAGCGACAGACACACATCATTTCTGGAGCCCTGATGATTAACAAGATGAGAA 1864
Oy 543 ThrHisAspSerGlu-----ValLeuLeuLysValMetAspSerArgHis 558
Db 1865 GAACTCTGAGAGAGCAAGATAAAGTGAATCTCTCAAGCTTGAAGCCCGACAGG 1924
Oy 559 AsnCysMetGluSerPheLeuGluAlaIleSerLeuMetSerGlnValSerTyrPro 578
Db 1925 GATATTTCCCTGCGCTCTTCCAGGACAGCCAGATGATGAGACAGCTCCCAACAAAC 1984
Oy 579 LeuValLeuLeuHisGlyValCysMetAlaGly---AspSerIleMetValGlnGlu 597
Db 1985 ATCGTGTACTCTATGCGTGTGTGCGGACGCTGAGATATCATATGATGAGAGATT 2044
Oy 598 ValTyrLeuGluAlaIleAspMetTyrLeuArgLysArgLysIleValSerAlaSer 617
Db 2045 GTGGAAGGGGTCTCTGATCTCTGATCTCTCATGACCGGAAAGCGATGTCCTACCA 2104
Oy 618 TrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAsp 637
Db 2105 TGAATATCAAAAGTTGCCAAACAGCTGCGCCCTGACCTCTGAGAGATTAAGAC 2164
Oy 638 LeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyLys 657
Db 2165 CTGCTCCATGCAATGTGTGACTAAAACCTCTCTCTGCGCGCTGAGGCGATGAGAGT 2224
Oy 658 Asn---ProPheIleLysLeuSerAspProGlyLysIleSerProThrValLeuSer 676
Db 2225 GAGTGTGCGCCGTCATCAACACTCATGACCGCCGATCCCATACGCTGCTCTAGG 2284
Oy 677 GluMetLeuThrAspArgIleProTyrValAlaProGlyLysGlnGluAlaGln 696
Db 2285 CAAGAATCATTTGAACGATCCATGATGCTCTGAGTGTGAGAGACTCCAAAGAC 2344
Oy 697 LeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGln 716
Db 2345 CTGACTGTGCGCTGACCAATGAGAGCTTGGAAACACGCTCTGGAATCTGTACAT 2404
Oy 717 GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGlnAsp 736
Db 2405 GCGCAGATCCCTTGAAGAACAGACGCTGATGAGAAAGAGATTTAAGAAAGCGG 2464
Oy 737 GlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMet 756
Db 2465 TGCAGGCCAGTACACCATCTGTAAGAGAGCTGCTACCTCATGACCCGCTGCATGAC 2524
Oy 757 TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeu 776
Db 2525 TATGACCCCATTCAGAGGCTTTCTCCAGCCATCATGAGACATTAATAATACCTGAA 2584
Oy 777 ThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAsp 796
Db 2585 GAGCAGAAATCCAGATATTGTTTCAGAAAAAACCACGACACTGAA----- 2629
Oy 797 LeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGlu 816
Db 2630 -----GTGAGCCCGCACACATTTTGAAGAGCGC 2656
Oy 817 HisLeuLysTyrIleSerLeuLeuGlyLysGlnAsnPheGlySerValGluLeuCys 836
Db 2657 TTCTTAAGAGAGATCCGATGTTGGAGAGGCGCACTTGGAGAGTTGAGCTCTGCGAG 2716
Oy 837 TyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln---His 855
Db 2717 TATGACCCCGAAGGAGCAATACAGGAGGAGGAGGCTGTTAAATCTCTAAGCCTGAG 2776
Oy 856 SerValProAspGlnIleArgAspPheGlnArgGluGlnIleLeuLysAlaLeuHis 875
Db 2777 AGTGAAGGTAAACCATGATGCTGATCTGAAAAAGAAATCGAGATCTTAAGAACCTCTAT 2836
```





QY 483 ValValArgArgGlyCysAsn-ProAlaProAlaProGlyCysSerProSerCys-CysA 502  
 Db 1917 -----CGAGGTGCTCTCCTGTCGACAGCAGCAGCATGTAATCCATGATATGTG 1967  
 QY 502 JaleuThrgInLeuSerPheHisThrIleProThrAsp 514  
 Db 1968 CTGAGACCCAGATGACCAAAACAGCGTG--ACTGAC 2002

## RESULT 6

US-09-962-854A-5  
 : Sequence 5, Application US/09962854A  
 : Publication No. US20030044803A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Pedersen, Finn S.  
 : APPLICANT: Annette, Soerensen B.  
 : APPLICANT: Hernandez, Javier Martin  
 : TITLE OF INVENTION: Methods for diagnosis and treatment of diseases associated with a  
 : TITLE OF INVENTION: expression of JAK1  
 : FILE REFERENCE: A-70020/RMS/DCF  
 : CURRENT APPLICATION NUMBER: US/09/962,854A  
 : PRIORITY FILING DATE: 2001-09-24  
 : PRIOR APPLICATION NUMBER: US 09/668,644  
 : PRIORITY FILING DATE: 2000-09-22  
 : NUMBER OF SEQ ID NOS: 5  
 : SOFTWARE: PatentIn version 3.1  
 : SEQ ID NO 5  
 : LENGTH: 975  
 : TYPE: DNA  
 : ORGANISM: Mus musculus  
 : FEATURE:  
 : NAME/KEY: misc\_feature  
 : LOCATION: (799)..(799)  
 : OTHER INFORMATION: "n" at position 799 can be any base.  
 US-09-962-854A-5

## Alignment Scores:

Pred. No.: 1,51e-41 Length: 975  
 Score: 483.50 Matches: 131  
 Percent Similarity: 61.62% Conservative: 52  
 Best Local Similarity: 44.11% Mismatches: 94  
 Query Match: 8.25% Indels: 23  
 Gaps: 9

US-09-397-967-16 (1-1099) x US-09-962-854A-5 (1-975)

QY 813 PheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerVal 832  
 Db 8 TTTGAAAAGCGTTCCTGAGAGAGATTCGTGACTTGGAGAGGCTCATTGGCAAGGTT 67  
 QY 833 GluLeuGlyArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGln 852  
 Db 68 GAGCTCTGACATATGATTCCTGAGGAGACAAACAGGAGGAGGTGCTGCAAGTCC 127  
 QY 853 LeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGlnIleGlnIleLeu 871  
 Db 128 CTGAAGCGTGAAGTGAAGTGAACATGATCTGAAGAGAGAGATCTTA 187  
 QY 872 LysAlaLeuHisSerAspPheIleValLysTyrArgGlyValIleSerTyrGlyProGlyArg 891  
 Db 188 CGGAACCTCTACATGAGAACATTTGTGAGTACAAAGGAATCTGCATGGAAGAGGAGGC 247  
 QY 892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGln 911  
 Db 248 AATGCTATCAAGCTCATGATGAGTTCCTGCGGAACGCTTAAGAGATATCTGCCA 307  
 QY 912 ArgHisArgGly---LeuHisThrAspArgLeuLeu-LeuPheAlaTrpGlnIleCysGly 930  
 Db 308 AAGATTAAGAACAAATCAACCTCAACAGCAAGCTAAATAATATGCAATTCAGAAATTGTA 367  
 QY 930 sGluMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIle 950  
 Db 368 GGGATGAGACTACTTGGGTTCTCGCAATTAAGTTCACCGGAGACTTACGACGACAG-AAATGT 426

QY 950 eLeuValGlnSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuP 970  
 Db 427 CCTTGTTGAGATGGATCCAGTATGATGATGAGACCTTGAGTTAACCAGCATTTG 486  
 QY 970 oLeuGlyLysAspTyrValValArgGluProGlyLeuInsProIleIleThrPTyrAl 990  
 Db 487 AAAGATTAAGAGTACTACAGATTCAGACACCGGGAAGAAAGCAAGTTCGGTACGC 546  
 QY 990 aProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyVal 1010  
 Db 547 TCCGGAATGTTTAATCCAGTGTATTTTAAAGCCCTCCGATGTCGGCTTGTGAGT 606  
 QY 1010 lValLeuTyrGlnLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLe 1030  
 Db 607 GACACTGCAGAGCTGCTCATTTACTGACTCCGAATTTAGTCCATGGCCTT----- 660  
 QY 1030 uArgHecMetGlyProGluArgGluGlyPro---ProLeuGlyArgLeuLeuGln---- 1047  
 Db 661 -----GTCGCCGAAAGGTAAAGCCCAACTCCAGCCAGAAACAAATTAAGGCGCT 710  
 QY 1048 -LeuLeuValGluGlyArgArg---LeuProProProProThrCysProThrGluValG 1066  
 Db 711 GTGCATCTACTAAGAGAGAGAGAGCCCTGCGATGTCACCAAA-TGCTGTGATGAAGTTAA 769  
 QY 1066 nGluLeuMetGlnLeuGlyCysTrpAlaProGlu-----ProHisAsp---ArgProAlaPh 1083  
 Db 770 CAGCTTATGGAGAAATTCCTGGAATTCGACTACTAACCAGAAATTTTCGAAACCTATG 829  
 QY 1083 eAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgProGly 1099  
 Db 830 GAAGAGTTTAAAGCCCC-----TTTAAATTAAGAGCTGCG 863

## RESULT 7

US-09-910-943-636  
 : Sequence 636, Application US/09910943  
 : Patent No. US20020081610A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Hemmati-Brianlou, Ali  
 : APPLICANT: Altman, Curtis  
 : TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression  
 : FILE REFERENCE: 7529/JG1480U1  
 : CURRENT APPLICATION NUMBER: US/09/910,943  
 : PRIORITY FILING DATE: 2001-07-23  
 : NUMBER OF SEQ ID NOS: 742  
 : SOFTWARE: PatentIn version 3.1  
 : SEQ ID NO 636  
 : LENGTH: 753  
 : TYPE: DNA  
 : ORGANISM: Xenopus laevis  
 : FEATURE:  
 : NAME/KEY: misc\_feature  
 : LOCATION: (1)..(763)  
 : OTHER INFORMATION: n may be a or g or c or t/u  
 US-09-910-943-636

## Alignment Scores:

Pred. No.: 1.31e-41 Length: 753  
 Score: 482.50 Matches: 102  
 Percent Similarity: 60.96% Conservative: 37  
 Best Local Similarity: 44.74% Mismatches: 71  
 Query Match: 8.23% Indels: 18  
 Gaps: 2

US-09-397-967-16 (1-1099) x US-09-910-943-636 (1-753)

QY 660 ProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeu 679  
 Db 76 CCATTCTAAAGAGCTTAGATGATCTGGAATACCAATTACTGTGTTAAACAACAAGACGG 135  
 QY 680 ThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuGlyLeu 699  
 Db 136 GTAGAGCCATTCATGATGATGCTCTGAAATGTGTTAGAGATTCAGAGATTAAGTGT 195





294 ILeValAspValSerIleAsnGlnAlaProArgVal-----GlyPro 307  
 1173 AA-STGGAGATTCAACTTACAGACGATGCCCCGGATCACTGACAGCTGCAGAGAACCC 1231  
 308 AlaGlyLinhIAsArgLeuValThrArgMetAspGlyHisIleLeu----- 324  
 1232 TTCCTGGCGGGGAGGATAGAGTACGACCAAGCGACAGCTGCTGTGCCAC 1291  
 325 -----GluAlaGluPheProGlyLeuProGluAlaLeu 335  
 1232 AAGGCAATTGTGGAGCCAGAGAACACACAGCTGAGTTC-----GAGTGCCC 1339  
 336 SerPheValAlaLeuValAspGlyTyrPheArgLeuIleCys----- 349  
 1340 CGCTTGCTTCTGGGACAGTGGTCTTGAG-----TGGCGTGTCCACATCTGGC 1393  
 350 -----AspSerArgHisTyrPheCysIys-----GluValAlaPro 361  
 1394 GGCACAGACAGCGCGCTTCAAGGTCATGTGAAGTCCCCCGCTGCCCTGCGA 1453  
 362 ProArgLeuLeuGluGluGluAla---AspValCysHisGlyProIleThrLeuAspPhe 380  
 1454 CCTCGGCTCTTACCAAGCAGAGCGCGCAGCTTGCTGTCCCGCTG----- 1501  
 381 AlaIleHisIysLeuLysAlaIleAspSerLeuPro---GlyThrTyrIleLeuArgArg 399  
 1502 -----CTCTCGTCTCTGGGAGTGGAGCCCATCTCCACTGTCCGCTGCACATAC 1549  
 400 SerProGlnAspTyrAspSerPheLeuThrAlaCys----- 412  
 1550 CGGCCAGACAGTACCTGGA-CTGCTGCACCATTTGTGGAGCCCGCAGTGAACGT 1608  
 413 ---ValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSer 431  
 1609 GAGGTTAATGAACCTGAGGCCAAGACAGATACAGTGTTCGTGTGACCTGAGCGGCC 1668  
 432 -----GlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeu 446  
 1669 ACGGAGAGAGAGAGGGGCG-----CTGGGGCGCTCCACCCCTCAT---GACCAAC 1716  
 447 ArgGluLeuLeuAlaIleCysThrPasnSerGlyLeuArgValAspGlyAlaIleLeuTyr 466  
 1717 ACAT-CTCCCTGAGCTTGTGTGACGCGCTGTGGAGGCGT---GGCATGTGGAGGCA 1772  
 467 LeuThrSerCysAlaIleProArgProLysGluLysSerAsnLeuIleValAlaArgArg 486  
 1773 CTGACC----- 1778  
 487 GlyCysAsnProAla---ProAlaProGlyCys-----SerProSer 499  
 1779 GGCTGGAGTGAAGCTGCTTGGCTGTGGTGGCGGCACTGCTGGGAGAGGTTTCC 1838  
 500 CysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluThrHis 519  
 1839 TCCTGGCGCTGT----- 1850  
 520 GluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgGluVal 539  
 1851 ---GGGACGGGACACGGGGCAG-----GAGCGGGGGGAGAACGTC 1888  
 540 AspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsn 559  
 1889 TCATCCCCCAGGCGCGCTGCTGCTGACGGGACTCAGCGCTGCACCCACTACAG 1948  
 560 CysMetGluSerPheLeuGluAlaIleAspLeuMetSerGlnValSerTyrPro---His 578  
 1949 ---CTGGAGTGTGAGCTTACACATGCACTTCCTGGCGCGCTGCGCCCTGACAG 2005  
 579 LeuValLeu-----LeuHisGlyValCysMetAla 588  
 2006 GTGCTTCTGCCCCCAGTGGGCTCCAGGCCCCGACACCTCCAGCCCAAGGCGCTTCTCA 2065

589 GlyAspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608  
 2066 GACTCCGAGATTCAGTGT----- 2083  
 609 LysArgGlyHisLeuValSerAlaSerThrLysLeuGlnValThrLysGlnLeuAlaTyr 628  
 2084 -----ACATGAGAG----- 2092  
 629 AlaLeuAsnTyrLeuGlnAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648  
 2093 -----CACCGGAGGCTGTGCT---GGCCCAATATCCAAAGTACGTTGTG 2134  
 649 LeuLeuAlaArgGlyGlyLysAspGlyAsnProProPheIleLysLeuSerAspProGly 668  
 2135 GAGGTGACAGTGGCTGGGGGTGCAGAGACCCAGCTGGATGACCTGGACAGCGCTGAG 2194  
 669 ValSerProThrValLeu----- 674  
 2195 GAGACAGACACCATCATCCGTGCTCAACGCCAGCAGCGCTACCTCTTCCGCATGCGG 2254  
 675 ---SerLeuGluMetLeuThrAspArgIleProThrValAlaProGlyCysLeuGlnGlu 693  
 2255 GCCAGCATTCAGGGCTCGGGGAC-----TGG----- 2281  
 694 AlaGlnThrLeuCysLeuGluAlaAspLysTyrGlyPheGlyAlaThrThrTrpGluVal 713  
 2282 ---ACCAACACAGTACAGAGTCCACCCGCGCAAGCGGCTGCAGGCTGAGGGCCCA 2335  
 714 PheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyr 733  
 2336 GTCCAAAGAGAGCGCGGCTGTGAAGGGCTGTGATCAGCAGCTGATCTCGGGCTGTG 2395  
 734 GluAspGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyIleThr----- 752  
 2396 -----GCTCCGTGTCTGCCACCTGCTCACCATCTGGCCCTTTTAAACCTG 2446  
 753 -----GlnCysMetAlaTyrAspProGlyArgArgProSerPhe----- 765  
 2447 GTGTGATCCGCGAGAGCTGCTG-----CATCGAGAGCGACCTTCACCTAC 2494  
 766 -----ArgAlaIleLeuAlaArgAspLeuAsnGlyLeuIleThrSerAsp 779  
 2495 CAGTCAGGCTCGGCGAGAGACCATCTGCACTTGCAGCTGACGAGCTTGCACCTTACC 2554  
 780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVal 799  
 2555 CGCGGCGCAAACTGCAGCGCCGAGCCC---CTGAGCTACCA----- 2593  
 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGlnArgHisLeuLys 819  
 2594 -----GTGCTAGAGTGGGAGGACATCCACC 2617  
 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerVal-----GluLeuCysArgTyr 837  
 2618 TTTCAGAGCTCTTCGGGAGGAACTTCGGGAGGTCACTCCGGGCAATGATCAAGAG 2677  
 838 AspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu---GlnHisSer 856  
 2678 GACGGGCTGAGATGTAAC-----GCAAGCATCAAAATGCTGAAGAGGTATGCC 2725  
 857 ValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeu---LysAlaLeuHis 875  
 2726 TCTGAATAATGACCATGTGACTTTCGGGAGAACTGGAAGTCTGTGCAAAATGGGGCAT 2785  
 876 SerAspPheIleValLysTyrArgGlyValAspTyrGlyProGlyArgGlnSerLeuArg 895  
 2786 CACCCCAACATCACTCAACCTCTGGGGGCTGT-----AAGAAGGAGGTATCTGTAT 2839  
 896 LeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly 915  
 2840 ATCGATATGAATATGCCCCCTAGCGGAACTCTGATTTTCTGGGAAACCCGGGTC 2899  
 916 LeuHisThrAsp----- 919



QY 494 roglYcysSerProSer-----CysCysAlaLeu-----ThrGlnLeuSerP 508  
 Db 1170 CCGGCTCCAGCCAGAGAGAGCTCCAGAGTGTGAGACTCTGAGAGAGATCAAGATTACCT 1229  
 QY 508 he-His-----ThrLeuPro----- 512  
 Db 1230 ATACATCTCAGCATGGCCGAGAGAGCTGCTGACCTCAGCTCTCCAGAACCTGCAAGT 1289  
 QY 513 -----ThrAspSerLeuGluTrp-----His 519  
 Db 1290 AATCCGGGAGACGAATCTGCAACAATGGCGCTACTGCTGACCTGACCTGCAAGGCTGGGCAT 1349  
 QY 520 GluAsn-LeuGlnHisGlySerPheThrLysIlePheArgGlySerArgArgGlyAlaVal 539  
 Db 1350 CACCTGGCTGGGCTGGCTGGCTGACAGGGAATG-----GGCAGTGGACTGGCCCTCAT 1403  
 QY 539 LAspGlyLysThrHisAspSerGluValLeuLysValMetAspSerArgHisArgAs 559  
 Db 1404 CCACCATACACCCACCTCTGCTTCCGTGCACACGGTCCCTGGGACCACTCTTTCGGAA 1463  
 QY 559 nCysMetGluSerPheLeuGlnAlaAlaSerLeuMetSerGlnValSerLysProHisLe 579  
 Db 1464 C-----CCGCACCA 1472  
 QY 579 uValLeuLeuHisGlyValAlcysMetAlaGlyAspSerIleMetValGlnGluPheValTy 599  
 Db 1473 AGCTCTGCTCCACACTGCCAACCCGCGCAGAGAC-----GAGTGTGTGGG 1517  
 QY \* 599 rLeuGlyAlaIleAspMetGlyLeuArgLysArgGlyHisLeuValSerAlaSerTrpLys 619  
 Db 1518 CGAGGGCTGGGCTGGCACCACAGCTGTGGCCCGAGGGCAC-----TGCCTGGG 1565  
 QY 619 sLeuGlnValThrLys-----GlnLeuAlaTyAlaLeuAsnTrpLeuG 634  
 Db 1566 TCCAGGGCCACCCACAGTGTGTCACTGACAGCCAGTTCCTCCGGGGCCAGAGTGGCTGA 1625  
 QY 634 uAsp-----LysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLe 650  
 Db 1626 GGAATCCCGAGTACTGACAGGGCTCCCGCAGGAGATGTGAATGCACGACACTGTTGGC 1685  
 QY 650 uAlaArg-----GlnGlyGlyAsp----- 656  
 Db 1686 GTGCCACCTGAGTGCAGCTCCACAGAAATGGCTACAGTACCTGTTTGGACCGGAGCTGA 1745  
 QY 657 -----GlyAsnProPhePheLeuLysLeuSerAspPro-- 667  
 Db 1746 CCAGTCTGTGGCTGTGCCACTATAGAGAGACCTCCCTTCTGCTGGCCCGCCCGCCAG 1805  
 QY 668 -GlyValSerProThrValLeuSerLeuGlnMetLeuThrAspArgIleProTrpValAl 687  
 Db 1806 CGGTGTGAACCTGACCTCTCTCTACATGCCCATC-----TGGCAAGTT 1847  
 QY 687 aProGlu-----CysLeuGlnGlnAlaGlnThrLeuCysLeuG 700  
 Db 1848 TCCAGATGAGAGGGGGGATGCCAGCCTTGGCCCATCAACTCACCACCTCTGTGTGGA 1907  
 QY 700 uAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheGlnArgGlyProAlaHis 720  
 Db 1908 CCTGTGTGACAGAGGCTGCCCGCC-----GAGCAGAGAGCCAGGCC 1949  
 QY 720 sIleThrSerLeuGluProAlaLysLeuLysPheTrpGluAspGlnGlyGlnLeuP 740  
 Db 1950 TCTGACGTCCATCTGCTGTGGCGTG-----GTTGGCATTTCTGCT 1988  
 QY 740 cAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyAspProG 760  
 Db 1989 GGTCTGTGCTTGGGGGTGCTTGGGATCTCTCATCAAG---CCACGGCAGCAAAAGAT 2045  
 QY 760 yArgArgProSerPheAlaGlnAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTy 780  
 Db 2046 CCGGAAGTACACAGATGCGGAGACTCTGCAGAAACG----- 2082  
 QY 780 rGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuGlyValAl 800

Db 2083 -GAGCTGTGGAGACCCCTGACACTAGCCGAGAGTCCCAACCA----- 2127  
 QY 800 aGlyAlaGlnLeuTyAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLysTy 820  
 Db 2128 ---GGCAGATG-----CGATCCGTGAAGAGACGGAGTGAAGAA 2165  
 QY 820 rIleSerLeuGlnGlyLysGlyAsnPheLysSerValGlnLeuGlyAspArgTyAspProLe 840  
 Db 2166 GGTGAAGTGTCTGGATGTGGCTTGTGGACAGCTCCAAAGGCAATGCTGATGCTCTGAA 2225  
 QY 840 uGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu-----GlnHisSerValProAs 859  
 Db 2226 TGGGAGAAATGTGAATTCSCAGTCCGCTCCAAAGTGTGAGAGAAACATCCSSCA 2285  
 QY 859 rGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheI 879  
 Db 2286 AGCCAAACAAAGAACTGTAGACGAAACATAGATGAGCTGGTGGTGGCTCCCAATGT 2345  
 QY 879 eValLysTyArgGlyValSerTyGlyProGluArgGlnSerLeuArgLeuValMetI 899  
 Db 2346 CTCCCGCTTGTGGCATCTGCTG-----ACATCCAGGTCGACAGTGTGACAA 2396  
 QY 899 uTyLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly---LeuHisTrh 918  
 Db 2397 GCTTATGCCCTATGGCTGCTCTTACAGCATGTCCGGAAACCGCGGACGCTGGGCTC 2456  
 QY 918 rAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGlyTyTrpLeuGlyAlaArg 938  
 Db 2457 CCAGGACCTGTGAATCTGTATGCGAGTATGCCAAAGGGAGTACAGCTGAGAGATGT 2516  
 QY 938 gArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGlnAlaHisVa 958  
 Db 2517 GCGCTCTGACACAGGAGACTTGGCGCTGGAAAGTGTGTGAAGAGTCCCAACCACT 2576  
 QY 958 lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlnLysAspTyTyTrpValVa 978  
 Db 2577 CAAATTTACAGACTTCGGGCTGGCTGCTGCTGGACATTCAGAGACAGACTACCAATGC 2636  
 QY 978 lArgGluProGlyGlnSerProIlePheTrpTyAlaProGluSerLeuAspAsnI 998  
 Db 2637 A---GATGGGGGCAAGTGGCCCATCAAGTGAATGGCGGTGAGTCCATCTCCCGCGCG 2693  
 QY 998 ePheSerArgGlnSerAspValTyTrpSerPheGlyValValLeuTyGlyLysLeuPheThrTy 1018  
 Db 2694 GTTCAACCCACAGAGTATGTGTGAGTATGTGTGACTGTGTGGAGCTGATACATTT 2753  
 QY 1018 rCysAspLysSerCysSerProSerAlaGlnPheLeuArgMetMetGlyProGluArgG 1038  
 Db 2754 TGGGGCCAAA-----CCTTACGA 2771  
 QY 1038 uGlyProProLeuCysArgLeuLeuGlnLeuLeuAlaGlnGlyArgGluProProP 1058  
 Db 2772 TGGGATCCAGCCCGGAGATCTCTGACCTGTGTGAAAAAGGGGAGCGCTGCCACCC 2831  
 QY 1058 oProThrCysProThrGluValGlnGlnLeuMetGlnLeuCysTrpAlaProGluProH 1078  
 Db 2832 CCCCATCTGCACCAATGTATGTCTACATGATGATGCTCAAAATGTTGATGATGATCTGA 2891  
 QY 1078 sAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgP 1098  
 Db 2892 ATGTCCGCGCAAGATTCGCGGAGTGTGTGTGAAT--TCTCCCGCATGCCAGGAGACCC 2949  
 QY 1098 o 1098  
 Db 2950 C 2950  
 RESULT 10  
 us-09-811-115-2  
 ; Sequence 2, Application us/09811115  
 ; Patent No. US20020035736A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Erickson, Sharon



```

Qy 687 aProglu-----CysLeuGlnGluAlaGlnThrLeuCysLeuG 700
Db 1848 TCCAGATGAGAGGCGGCATCCACCCCTTGCCTCCATCACTGCACCCACCTCTGTGTGA 1907
Qy 700 uAlaAspLysTrpGlyPheGlyAlaThrThrTropIuValPheGlnArgGlyProAlaH 720
Db 1908 CCTGGATGACACAGGCTGCCCCGCC-----GAGCGAGAGACCCAGGCC 1949
Qy 720 sIleThrSerLeuGluProAlaLysLysLeuLysPheThrGlnLysArgGlnGlyLeuPr 740
Db 1950 TCTGACGTCCATCGTCTGGGGTG-----GTTGGCATCTCTCT 1988
Qy 740 cAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProG 760
Db 1989 GGTGCTGTCTTGGGGGCTGCTTGGGATCCTCATCAAG---CGACGCGACAGACAGAT 2045
Qy 760 yArgArgProSerPhePheAlaIleLeuArgAspLeuAsnGlnGlyLeuIleThrSerAsp 780
Db 2046 CCGGAAGACACGATCGGAGACCTGCTCAGGAAACG----- 2082
Qy 780 rGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAl 800
Db 2083 -GAGCTGTGGAGCGCGTGCACACCTAGGAGGATGCCCAACG----- 2127
Qy 800 aGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLysT 820
Db 2128 ---GCCGAGATG-----CGCATCCTGAAGAGACGAGCTGAGGA 2165
Qy 820 rIleSerLeuGlnGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProL 840
Db 2166 GGTGAAGGTCTGTGATGCTGCGCTTGTGGCACAGTCTACAAAGGCATCGATCGATCTGA 2225
Qy 840 uGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu---GlnHisSerValProAs 859
Db 2226 TGGGAGAGATGTGAATAATTCAGTGGCCATCAAGTGTGAGGGGAAACACATCCCCCA 2285
Qy 859 pGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheI 879
Db 2286 AGCCAAACAAAGAACTTTCAGCAAGACATACCTGATGCTGCTGCGCTCCCATATGT 2345
Qy 879 eValLysTyrArgLysValSerTyrGlyProGlyArgGlnSerLeuArgValMetG 899
Db 2346 CTCCCGCTTGTGGCATCTGCTG-----ACATCCAGGCTGCTGATACACA 2396
Qy 899 uTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly---LeuHisTh 918
Db 2397 GCTATGCGCCTATGAGCTGCTGCTTATGACCAATGTCGCGGAAACCCGAGCGCTGGCTC 2456
Qy 918 rAspArgLeuLeuLeuPheAlaIleTrpGlnIleCysLysGlyMetGluTyrLeuGlyAla 938
Db 2457 CCAGGACCTGCTGAACCTGCTATGCTATGCCAAGGGGATGAGCTGCTGAGAGATGT 2516
Qy 938 gArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHisVa 958
Db 2517 GCGGCTGTACACAGGAGCTTGGCGCTCGAGACCTGCTGTCAGAGATGCCAACCAATGT 2576
Qy 958 lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlnLysAspTyrTyrValVa 978
Db 2577 CAATAATTACAGACTTGGGCTGCGCTGCTGACATTCGACGACAGAGATACCAGTGC 2636
Qy 978 lArgGluProGlyLysSerProIlePheTrpTyrAlaIleProGlnSerLeuSerAspAs 998
Db 2637 A---GATCGGGGCAAGGTGCCATCAAGTGAATGAGTCCATTCATCCGCGCGCG 2693
Qy 998 ePheSerArgLysIleAspValIleTrpSerPheGlyValValLeuTyrGluLeuPheThrTy 1018
Db 2694 GTTACACCCACAGAGATGATGTGTGTGAGTATGTGTGCTGTGTGGAGCTGATGACTTT 2753
Qy 1018 rCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgG 1038
Db 2754 TGGGGCCAAA-----CCTTACGA 2771
Qy 1038 uGlyProProLeuCysArgLeuLeuGluLeuLeuAlaGluLysArgLeuProProPr 1058

```

```

Db 2772 TGGGATCCACAGCCCGGAGATCCCTGACCTGCTGGAAGAGGGGAGCGGCTGCCAGCC 2831
Qy 1058 oProThrCysProThrGlnValGlnGluLeuMetGlnLeuCysTrpAlaProGluProH 1078
Db 2832 CCCCATCTGCACCATGATGATGCTTACATGATCATGTGTCAAAATTTGGATGATGACTTGA 2891
Qy 1078 sAspArgProAlaPheAlaIleThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgPr 1098
Db 2892 ATGTGCGCCAGATTCCTCGGAGATTGTGTGTGAAT--TCTCCCGCATGGCCAGGAGCC 2949
Qy 1098 o 1098
Db 2950 C 2950

RESULT 11
US-09-877-177-11
; Sequence 11, Application US/09877177
; Publication No. US20020192652A1
; GENERAL INFORMATION:
; APPLICANT: Peter V. Danenberg et al.
; TITLE OF INVENTION: Method of determining Epidermal Growth
; FACTOR RECEPTOR: Factor Receptor and HER2-New Gene Expression
; FILE REFERENCE: 11220/120
; CURRENT FILING DATE: US/09/877,177
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177-11

Alignment Scores:
Pred. No.: 1,2e-39 Length: 4530
Score: 475.00 Matches: 261
Percent Similarity: 36.85% Conservative: 116
Best Local Similarity: 25.51% Mismatches: 337
Query Match: 8.11% Indels: 312
DB: 9 Gaps: 46

US-09-397-967-16 (1-1099) x US-09-877-177-11 (1-4530)
Qy 223 ProCysGlyArg---LeuProGlyArgProTyr---AlaLeuMetAlaLysTyrIleLeu 240
Db 531 CCCTGTACAGAGGCGCTCCCGAGAGGCTGCGGAGGCTGCAAGCTTCGAAGCTTCACAGA 590
Qy 241 AspLeuGluArg-----LeuHisProAlaAlaIleThrGluThrPheArgValGlyLeu 258
Db 591 GATCTTTGAAGAGGAGGCTTGTATCCAGCG----- 620
Qy 259 ProGlyAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyLeuProTrp 278
Db 621 -----GAAACCCCAAGCTTGTCTA-----CCAGGA 644
Qy 279 SerSerAsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspValSer 298
Db 645 CAC-----GATTTTGTGGA----- 659
Qy 299 lLeaGlnAlaIleProArgValGlyProAlaGlyLysHisArgLeuValThrValThrArg 318
Db 660 ---GGACATCTTCCACAGAACACACACCTGCTCTAC----- 695
Qy 319 MetAspGlyHis----- 322
Db 696 ACTGATACACACCAACCGCTCTCGGCGCTGCCACCCCTGTTCTCCGATGTGAAGGCTC 755
Qy 323 ---lLeuGlnAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuVal 341
Db 756 CCGCTGCTGGGAGAGAGATTTCTGAGGATTTCTGAGAGC--CTGAGC----- 798
Qy 342 AspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaPro 361

```

Db 799 -----GCACACTGCTGTGGCGGTGGCTGTGCGCCGTCAAGGGGCGACTGCC 846  
 QY 362 ProArgLeuLeuGluGluGlu-----  
 Db 847 ACTGACTGTGTCATGAGAGGTGTGCTGGCGGTGCACGGGGCCCAAGACTGTACTGC 906  
 QY 369 -----AlaAspValCys-----HisGlyProIleThrLeu 378  
 Db 907 CTGGCGCTCCACTTCACACACAGTGGCATGTGTGAGCTGCTGCCAGCCCTGTGC 966  
 QY 379 AspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrlleLeuArg 398  
 Db 967 ACCTACACACAGACACAGCTTTGAGTCCATCCCAATCCGAGGGCCGGTAT----- 1017  
 QY 399 ArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaLysValGlnThrProLeuGly 418  
 Db 1018 -----ACATGGCGCGCGCTGTGACTGCTGTGCTTCCCTTCACAACTACTTTCT 1065  
 QY 419 ProAspTyrLysGlyLysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGly 438  
 Db 1066 ACGGACGTGGATTCCTGCACCT--CGTGTGCGCCCT----- 1100  
 QY 439 LeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaLysTrpAsn--SerGlyL 458  
 Db 1101 ---GCACACACAGAGGTGACAGCAGAGA-----TGACACACAGCGGTG 1142  
 QY 458 euArg-----ValAspGlyAlaAlaLeuTyrLeuThrSerCysCys- 471  
 Db 1143 TGAGAGGTGACAGAGCGCTGTGCGCGAGTGTGTATGTGTGGCGATGGAGCACTTGGC 1202  
 QY 472 -----AlaProArgProLysGlySerAsnLeuIleValValArgArg- 486  
 Db 1203 AGAGGTGAGGCGAGTTACACAGTGCATATCCAGAGCT--TTGCTGTGTCGACAGAGAT 1259  
 QY 487 -----GlyCysAspProAlaProAla 494  
 Db 1260 CTTTGGAGCGCTGGCATTTCTGCCGAGAGCTTTGATGGGAGCCAGCCTTCACAACTGC 1319  
 QY 494 roGlyCysSerProSer-----CysCysAlaLeu-----ThrGlnLeuSerP 508  
 Db 1320 CCGGCTCCAGCCAGAGAGCTCCAGAGTGTGTGAGACTGTGAGAGAGATACAGAGTTACT 1379  
 QY 508 he-His-----ThrIlePro----- 512  
 Db 1380 ATACATCTCAGCATGGCCGAGACAGCCCTGCAGCTCAGCGCTTCCAGAACCTGCAAGT 1439  
 QY 513 -----ThrAspSerLeuGluTrp-----His 519  
 Db 1440 AATCCGGGGAGCAATTTTGCACAAATGGCGCTACTCGCTGACCTGCAAGGCGTGGGCAT 1499  
 QY 520 GlnAsn--LeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluVala 539  
 Db 1500 CAGCTGCTGGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1553  
 QY 539 LaspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAs 559  
 Db 1554 CCACCATTAACACCACTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1613  
 QY 559 ncysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLe 579  
 Db 1614 C-----CGCGACCA 1622  
 QY 579 uValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTy 599  
 Db 1623 AGCTGTGCTCCACACATGCCCAACGGCCAGAGAC-----GAGTGTGTGGG 1667  
 QY 599 rLeuGlyAlaAlaIleAspMetLysLeuArgGlyArgGlyHisLeuValSerAlaSerTrpLy 619  
 Db 1668 CGAGGGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1715  
 QY 619 sLeuGlnValThrLys-----GlnLeuAlaTyrAlaLeuAsnTyrTrpLeuG 634

Db 1716 TCCAGGGGCCACCCAGTGTGTCAACTGCAGCCAGCTTCTTGGGGCCAGAGTGTGTGA 1775  
 QY 634 uAsp-----LysGlyLeuProHisGlyAsnValSerAlaArgLysValLeu 650  
 Db 1776 GGAATGCCGAGTACTGACGAGGGGCTCCAGGAGACTATGTATGATCCAGGACTGTGGC 1835  
 QY 650 uAlaArg-----GluGlyGlyAsp----- 656  
 Db 1836 GTGCCACCTGAGTGTGACGCCCGAGAAATGGCTGATGACCTGTTTGGACCGGAGGTGA 1895  
 QY 657 -----GlyAsnProPheIleLysLeuSerAspPro-- 667  
 Db 1896 CCAGTGTGTGCGCTGTGCCACTATTAAGACCTCCCTTGTGCGGTGGCCCGCTCCAG 1955  
 QY 668 -GlyValSerProThrValLeuSerLeuGlnMetLeuThrAspArgLysProThrValAl 687  
 Db 1956 CGGTGTAAACCTACCTCTCTCTACATGCCATC-----TGGAAGTT 1997  
 QY 687 aProGlu-----CysLeuGlnGluAlaGlnThrLeuCysLeuG 700  
 Db 1998 TCCAGATGAGGAGGCGCCATGCCAGCTTGGCCCATGACAGTGCACCACTCTGTGTGA 2057  
 QY 700 uAlaAspLysTrpLysPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaI 720  
 Db 2058 CTTGGATGACAGGGGCTGCCCGCC-----GAGCAGAGAGCCAGGCC 2099  
 QY 720 sIleThrSerLeuLeuProAlaLysLysLeuLysPheTyrluAspGlnGlyGlnLeuPr 740  
 Db 2100 TCTGACCTCCATCTCTCTGTGGGTG-----GTGGCATTTCTGCT 2138  
 QY 740 cAlaLeuLysTrpThrGluLeuLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProG 760  
 Db 2139 GGTGCTGTCTTGGGGGTGTGCTTGTGGATTCCTATGACAG--CGACGGCAGCAGAAAT 2195  
 QY 760 YArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTy 780  
 Db 2196 CCGGAGTACACAGATGGGAGACTGCTGCAGAAAG----- 2232  
 QY 780 rGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAl 800  
 Db 2233 -GAGCTGTGTGAGCGGCTGACACCTACCGAGCGATGCCCAACAG----- 2277  
 QY 800 aglyAlaGlnLeuTyrLysGlnAspProAlaIlePheGlnGluArgHisGlyLeuTy 820  
 Db 2278 ---GCCCAGATG-----CGGATCTCGAAGAGAGCGAGCTGAGAGAA 2315  
 QY 820 rIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLe 840  
 Db 2316 GGTAAAGTGTGTGATGTGGCTTTTGGCAGCTGACAGGCGCATGTGATCCCTGA 2375  
 QY 840 uGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu--GlnHisSerValProAs 859  
 Db 2376 TGGGGAAGATGTGAATAATTCAGATCAAGATGTGAGGAGAAACATATCCCAAA 2435  
 QY 859 pGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheI 879  
 Db 2436 AGCCAAACAAGAAATCTTACAGAGCATACGTATGCTGTGGCTGCCCATATGT 2495  
 QY 879 eValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetG 899  
 Db 2496 CTCCCGCTTCTGGCATGTGCTG-----ACATCCAGGTGACGTGTGAGACACA 2546  
 QY 899 uTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly--LeuHisTh 918  
 Db 2547 GCTTATGCCCTATGCTGCTCTTACACATGTCCGGGAAACCGGAGCGCTGGGCTC 2606  
 QY 918 rAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaAr 938  
 Db 2607 CCAGGACCTCTGTAACAGGTATGACAGATTGCCAAGGGGTAGAGCTACCTGAGAGATGT 2666  
 QY 938 gArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVa 958  
 Db 2667 GCGGCTGTACACAGGAGACTTGGCGGCTCGGAAAGTGTGTGCAAGAGTCCCAACACATGT 2726

Oy	958	ILYSILIALIASpPheGlyLeuNAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVa	978
Db	2727	CAAAATTACAGACTTCGGGCTGGCTGGCTGGACATTGACAGACAGATTACCATTGC	27866
Oy	978	LARGluProGlyGlnSerProIlePhePrrTyrAlaProGlnSerLeuSerAspAsnI	998
		:::    :::            :::	
Db	2787	A---GATGGGGGCANAAGTGGCCATCAAGTGGATGGCGGTGGAGTCCATTCCTCCGGCGG	28438
Oy	998	ePheSerArgGlnSerAspValTrrPserPheGlyValValLeuTyrGlnLeuPheThrTy	1018
Db	2844	GTTTACCACACAGAGTATGTGTGGATTATGGTGTGAGCTGTGTGGACCTATGACTTT	29030
Oy	1018	rCyAspIlySerCysSerProSerAlaGluPheLeuArgMetuGlyProGluIargGl	10388
		:::	:::
Db	2904	TGGGGCCCAAA-----CCTTACGA	29212
Oy	1038	uGlyProProLeuGlnCysArgLeuGlnLeuGlnLeuAlaGlnGlyArgGluProProP	10586
Db	2932	TGGGATCCCAAGCCCGGAGGATCCCTCAACCTGCTGGAAAGGGGGAGCGGCTCCCAACG	29818
Oy	1058	oProThrCysProThrGluValGlnGluLeuMetGlnLeuGlyCysTrpAlaProGluProH	1078
		::::	
Db	2982	CCCCATCTGCACCACTTCATGCTCTACATGATGATGCTCAATGTTGGATGATTGACTCTGA	30411
Oy	1078	sAspArgProAlaPheAlaThrLeuSerProGlnLeuAspArgProLeuTyrArgGlyArgP	10988
		:::	
Db	3042	ATGTCGGCCAGATTTCCGGGAGTTGGTGTGTAAT--TCTCCCGCATGGCCAGGGAGCCG	30999
Oy	1098	o 1098	
Db	3100	C 3100	

```

RESULT 12
US-09-811-123-7
; Sequence 7, Application US/09811123
; Patent No. US2002001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; FILE REFERENCE: GENEENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/338,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 9274
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Vector Sequence
US-09-811-123-7

Alignment Scores:
Pred. NO: 3.5e-39 Length: 9274
Score: 475.00 Matches: 316
Percent Similarity: 36.09% Conservative: 134
Best local Similarity: 25.34% Mismatches: 419
Query Match: 8.11% Indels: 383
DB: 10 Gaps: 60

US-09-397-967-16 (1-1099) x US-09-811-123-7 (1-9274)
QY 9 Prolenite---ProGlnArgSerCysSerLeuSerSerSerClnUlaGlyAlaLeuHts 27
DB 1621 CCACGACGAGTCGAAAGAAATTCGAGGCGCGCCGCCGCCGCC---CCA 16777

```

QY	28	valleuueurproproadagglurproglurprotoprogluathleu--serpheserphiegluys	47
Db	1678	ccscctgcagacaccccgccggccggccgctccacccgggctccacccggagacatga--	17355
QY	47	ptylleuallagluaspreucysvalargalaalalysalacysgluyleleuprovaltu	67
Db	1736	-----gcggcgccgctttgcccsc--tg	1756
QY	67	rhissertleuhalaleuhalathrgluasrpheserystrrperproserhisl	87
Db	1757	ggcgctctctccgcccctc-----tgcccccscga----	1787
QY	87	eheueaslleuaspvalasprhringvalleuvaltutraglualeuarphethyrphert	107
Db	1788	-----gcccgacaccccaagt-----	1805
QY	107	casprtrphegluenglutrhrcysnhisarphiegluethgluyspleu--thse	126
Db	1806	-----tgacacggcagacagacatgaagctcgccgctcccgccag	1843
QY	126	rlalaleuaspheuhistalleugluhlsleuphealaglnhlsatgseraspleuval	146
Db	1844	tcgccagacccacctgcagatgctcgccacaccttac-----cacggctccacagctgt	1897
QY	146	lsertilyarg-----leuprovallguyleusermetuylugluingyl	161
Db	1898	gcagsgaactgcgaactacactacactgcacacacatggccacacctg-----tc	1945
QY	161	upheleuserleualavalaleuaspheualaglnmetvalatrgluglnalaglnatg--	180
Db	1946	cttcctgcag-----gatattccagaggt--gcagggctacagctgcctacatgcctc	1992
QY	181	progluylleuileuysttrhvalasertyrlasacyslleuproposerleuar	199
Db	1993	acacacagtgagggcagctccacactgcagagcctggcgatgt-----gcagsgca	2043
QY	199	gasprvallleaglinsglasnbhevaltharagaratgleatgatgthvalvalle	219
Db	2044	cccgactcttgtagacgacacactatgccctgcctgcctagacacatggagaccccgctga--	2101
QY	219	ualaleuueuprocygluargt---leuproglyuarprotyr---alaleumetaleu	237
Db	2102	-cattaccacacccctgctacacagggcctcccccagagagcctggcgagactgcacgttcgaa	2160
QY	237	stytllleuaspheugluargt-----leuhisproalaalathrthrglutrhphear	255
Db	2161	gccttcacagagatcttgaaaggaggggtcttgattccaggg-----	2200
QY	255	gvalaglyleuproglyualaglnugluproglyleuualargvalalaglyaspaingl	275
Db	2201	-----cagcccccgactctgctga-----	2218
QY	275	ylleprotrpseraserasnapgluueupheglnthrphesaysapheprogluileua	295
Db	2219	-----ccagagacac-----gattttgtggaa-----	2239
QY	295	lasprvalserlleasnglnalaepraargvalaglyproalaglyluhlsatrgleuvalth	315
Db	2240	-----ggacatctttccacacagacacacacacacagctgcctgcac-----	2257
QY	315	rvalthrargmetasprgluylhs-----	322
Db	2276	actgatagacacacacacccgcttcggcgctgcacaccccgcttctccgatgt	2325
QY	333	-----lleuenglualagluupheprogluyleuuprogluualaleuserphava	338
Db	2326	gtaaaggctcccgctgctgggagagagatgtctgagagattgtcagagc--ctgagc--	2378
QY	338	lalaleuualaspolytyrtheargleuilecysaspserseralughisrtyrphesyslygl	358
Db	2379	-----gcgacatgcttttgcggctgcctgcctgcacccgctgcgaagg	2416

QY	181	----	ProGlyIuLeuLeuLeuYshPhaIaSerTrpYLaCyGLeuProProSerLeuAr	199
Db	1993	ACAAACCAAGTGAAGCAGGCTCCACCTGACGAGGCTGGCATGT	-----GCGAGCA	2043
QY	199	gAsPvaIIleGInGlyInsnPheVaIThrArGaRarGILeArArgThVaIaLe	219	
Db	2044	CCCAAGCTCTTGAGGACAACTATGCCCTGGCGCTGTATACAAATGAGACCCGGTGA	-	2101
QY	219	uAlaLeuLeuProCysGlyArg---	LeuProGlyArpProTyR--	AlaLeuMeTaLaLy
Db	2102	-CAATACCAACCCCTGTGACAGGGGGCCCTCCCAAGAGAGCCGTGGGAGAGCTGCAGCTTCGAA	2160	
QY	237	sTyRrIleuLeuSpheuLnuArg-----	LeuHISProAlaIaIaThrInGluThrPheAr	255
Db	2161	GCCCTACAGAGATCTTGAAGAGGAGGGGCTGTGATCCAGG	-----	2200
QY	255	gVaIGlyLeuProGlyLaGInGluInuProGlyLeuLeuArGvaIaIaGlyAspAnG	275	
Db	2201	-----GAACCCCAAGCTGCTGA	-----	2218
QY	275	yIleProTrpSerSerAsnAspGluLeuPheGInThrPheCysAspPheProGluIleVa	295	
Db	2219	-----CGAGGACAC	-----GATTTGTGTGA	2239
QY	235	lAsPvaISerIleAsnGlnAlaPArCaRvLaIGlyProAlaGlyGluHISArgLeuVaITh	315	
Db	2240	-----GGACATCTTCCACAGAACAAACAGCAAGCTGGCTGTAC	-----	2275
QY	315	rVaIThrArgMeTAspGlyHIS	-----	322
Db	2276	-----ACTATGAGACACCAACCGCTGTGGGCTGTGCAACCCCTTGTCCGATGT	2325	
QY	333	-----IleLeuGlnuAlaGluPheProGlyLeuProGluuAlaLeuSerPheVa	338	
Db	2236	GTAAGGCTCCCGCTGCTGGGAGAGAGTCTGAGGATGTGCAGAGC	-CTGAGC-----	2378
QY	338	lAlaLeuVaIAspGlyTrpPheArGleuIleCySAspSerArGhISrTrpPheCySlyG	358	
Db	2379	-----GCGACGTGCTGTGGCGGTGGCTGTGCCCGGTGAAGG	2416	

OY	358	uValAlaProProkArgLeuLeuGluGluGlu	-----AlaAspValCys-----HisGlyPro	375	366
Db	2417	GCCACCTGCCACCTGACTGTCTGCATGAGCAGTGTCTGCCGGCTGCACGGGCCCAAGCA		2476	
OY	369	-----	-----AlaAspValCys-----HisGlyPro	375	
Db	2477	CTGTGACTGCGCTGGCTGCCTCCACTCACTTCAACCCACGTGCATCTGTGAGGTCTGACCTGCC		2536	
OY	375	oileThrLeuAspPheAlaIleHisIstyLeuAlaAlaGlySerLeuProGlyIthThy		395	
Db	2537	AGCCCTGGTCACTCAGCAACACAGACACAGCTTTGATGTCATGCCCAATCCCGAGGGCCGGA		2596	
OY	395	rIleLeuArgArgSerProGlnIAspIAspSerPheLeuLeuThrAlaCysValGlnTh		415	
Db	2597	T-----	-----ACATTCGGCGCCACCTGTGACTGCTGTCCCTTCCATCA	2635	
OY	415	rProLeuGlyProAspPtyrIstyIstyLeuIleArgGlnAspProSerGlyAlaIlePhe		435	
Db	2636	CTACCTTTTCAGCAGCAGCTGGAGTCTGCACCCCT-CCCTGCCCTCT		2680	
OY	435	rIleuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuAlaIleCysTyrAs		455	
Db	2681	-----GCCACACCAAGAGTGCACAGCAGGA		2712	
OY	455	n-SerGlyLeuArg-----	-----ValAspGlyAlaAlaLeuIthLeuThr	468	
Db	2713	CACAGCCGTGTGAGAAATGTGCAGACAGCCCTGTGCCGATGTGCTATGTGTGGCAGTGG		2772	
OY	469	SerCysCys-----	-----AlaProArgProGlyIstySerAspIleuVal	483	
Db	2773	AGCACTTGCCAGAGGTGAGGGCAGATTACAGTCCAGTATCCAGGAGT---		2829	
OY	484	ValArgArg-----	-----GlyCysAsnPro	490	
Db	2830	GCAAGAAAGATCTTTGGAGGCTGCATTTCTGCCGAGAGCTTTGATGAGGAGCCAGCCT		2889	
OY	491	AlaProAlaProGlyCysSerProSer-----	-----CysCysAlaLeu-----Thr	504	
Db	2890	CCAACTGCCCCCTCCAGCCAGCAGACAGCTCCAACTGTTTGAAGCTGTGGAAGAGATCA		2949	
OY	505	GlnLeuSerPhe-His-----	-----ThrIlePro-----	512	
Db	2950	CAGTTTACCTATACATCTCAGCATGTGGCCGAGACCGCTGCCTACCTCAGCGCTTCCAGA		3009	
OY	513	-----	-----ThrAspSerLeuGluIthP-----	518	
Db	3010	ACCTGCAGATTAATCCGGGGAGCAATTTCTGCACATAGTCCGCTTACGCTGACCCCTGCAG		3069	
OY	519	-----HisGlnAsn-LenGlnHisGlySerPheThrIstyIlePheArgGlySerArg		536	
Db	3070	GCGTGGGCATCAGCTGGCTGGGGCTGGCGCTCACTAGAGGAATG-----	GGCACTGGAC	3123	
OY	536	rGluGluValAlaSerGlyGluIthHisAspSerGluValLeuLeuIstyValMetAspSer		556	
Db	3124	TGGCGCTCATCCACATTAACACCCACCTGCTGTCTGTGCACACGGTGCCTTGCGCCAGC		3183	
OY	556	rGlnIstArgAsnCysMetGluSerPheLeuGlnAlaAlaSerLeuMetSerGlnAlaSerT		576	
Db	3184	TCTTTCCGAAAC-----	-----	3194	
OY	576	yrrProHisLeuValLeuLeuHisGlyValAlaCysMetAlaGlyAspSerIleMetValGlnG		596	
Db	3195	---CCGACACCAAGCTGTCTGCTCCACATGCCACACCGGCCAGAGAGAC-----	---G	3237	
OY	596	IuPheValIthLeuGluAlaAlaIleAspMetIthLeuArgIstyArgIstyHisLeuValSer		616	
Db	3238	AGTGTGTGGGAGAGCGGCTGGCTGCACACAGCTGTGGCCCGAGAGGACAC-----	3287		
OY	616	IAspTrpPtyrIstyLeuGlnValIthThrIsty-----	-----GlnLeuAlaTyrAlaLeu	631	
Db	3288	---TGTGTGGGGTCCAGAGGCCACCCACTGTGTCAACTGCAGCCAGCTTCTGTGGGGCCACGG	3345		
OY	631	snTyrLeuGlnAsp-----	-----LysGlyLeuProHisGlyIstyValSerAlaArgI	647	

Db	3346	AGTGGCGTGGAGGAATGCCAGTACTGACGAGGCGTCCCGACGAGAGTATGTGAATGCCAAGC				3405
Oy	647	ysValLeuAlaIArg-----		gluIyglYAsp-----		656
Db	3406	ACTGTGTGCCGTGGCCACCTGAGTGTACAGCCCGAAGTGGCTCAGTACCTGTTTGGAC				3465
Oy	657	-----		glYsnPrpOpheLeuLysLeuS		665
Db	3466	CGAAGCTGACCACTGTGTGGCTGTGCCACTATAGAACCTCTCTTCTGCGTGGCC				3525
Oy	665	erAspPro--GlyValSerProThyValLeuSerLeuLueLueThLeuThrAspArgIleP				684
Db	3526	GCTGCCCCAGGCGTGTGAACCTGACCTCTCTTCTACATGGCCACT				3569
Oy	664	roTPrValAlaProGlu-----		cysLeuInGluAlaInPrL		697
Db	3570	--TGAAGTTCACAGATGAGAGGCGGCATGCCAGCCTTGCCCATCACTGACCCACT				3627
Oy	697	eucYsLeuGluAlaIAspLysTrpGlyPheGlyAlaThrTrpGluValPheGlnArg				717
Db	3628	CCTGTGTGACCTGTGATGACAGAGGCGTCCCGCCG				3669
Oy	717	lyProAlaHisLleThrSerLeuGluProAlaLysLysLeuLysPheTrpGluAspGln				737
Db	3670	GAGCGACCCCTGACCTCATCATGCTGTCTGGGATG				3708
Oy	727	lyGlnLeuProAlaLeuLysTrpTrpGluLeuAlaGlyLeuLleThrGlnCysMetAlaTr				757
Db	3709	GCATTCTGTCTGTGTGTCTTGGGGGTGTCTTTTGGATCTCTATCAAG				3765
Oy	757	yrAspProGlyArgArgProSerPheArgAlaLleLeuArgAspLeuSngLysLleTr				777
Db	3766	AGCAGAGAGATCCGGAATACACATGCGGAGAGCTGTGCGAGGAACG				3812
Oy	777	hrSerAspTrpGluLeuLeuSerAspProThrProGlyLleProSerProAspGluL				797
Db	3813	-----GACTGTGTGGAGCCGCTGACACTTACGCGAGAGCATGCCAACAG				3857
Oy	797	eucYsValAlaGlyAlaGlnLeuTyAlaCysGlnAspProAlaIlePheGlnGluArgH				817
Db	3858	-----GCCACATG				3885
Oy	817	IsLeuLysTyrlIleSerLeuLeuGlyLysIAsnPheLysSerValGluLeuCysAlaTr				837
Db	3886	AGCTGAGAGAGGTGAAGGCTTGGATCTGGCGCTTTGGCAGGTACAAAGGCACT				3945
Oy	837	yrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu--GlnHis				856
Db	3946	GGATCCCGTGGAGGGAATGTGAATAATTCATGTGCCATCAAAAGTGTGAGCGAANA				4005
Oy	856	erValProAspGlnGlnArgAspPheGlnArgGluLleGlnLeuLysAlaLeuHis				876
Db	4006	CATGCCCCCAAGGCCAACAAAGAAATCTTGAACACAAACATCACTGATGGTGTGGGCT				4065
Oy	876	erAspPheLleValLysTyArgIValSerTrpGlyProGlyArgGlnSerLeuArgL				896
Db	4117	TGTGTACACAGCTTAATCCCTATGCGTGTCTTGAACCATGTCCGGGAAACCCCGGAC				4176
Oy	916	--LeuHisLThrAspArgLeuLeuLeuPheAlaTrpIlnLleCysLysGlyMetGluTyrl				935
Db	4177	GCTTGGGCTCCACGACCTCTCTAAGTGGGTGTGAGATTGCCAAAGGGATGAAGTACC				4236
Oy	935	eucGlyAlaArgArgCysValHisArgAspLeuAlaIaIaArgsnLleLeuValIeUsrG				955
Db	4237	TGAGAGATGTGGGCTGTGACACAGGAGCTTGGCGCGCTGGAACGGTGGTCAAGACT				4266
Oy	955	IuaIaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspT				975







QY 995 erAspAnuIllePheSerArgInSerAspValTPrSerPheGlyValValLeuTyrGlu 1015  
 Db 4414 TCCCGCCGGCGGTTCACCCACAGAGTGTGTGAGTATGACTGTGGAGC 4473  
 QY 1015 euphethrTyrCysAspLysSerSerProSerAlaGluPheLeuArgMetGlyP 1035  
 Db 4474 TGATGACTTTTGGGCCAAA----- 4493  
 QY 1035 roGluAArgGluGlyProProLeuCysArgLeuGluLeuAlaGluGlyArgT 1055  
 Db 4494 --CCTTACGATGGGATCCACCCCGGAGATCCCTGACCTCTGGAAAGGGAGCGC 4551  
 QY 1055 eupProProProthCysProthGluValGlnGluLeuMetGlnLeuCysTrpAlaP 1075  
 Db 4552 TGCCCGACGCCCATCTGCACCATTTGATGTCTACATGATGATGTCATAATGTTGATGA 4611  
 QY 1075 roGluProHIsAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTPrA 1095  
 Db 4612 TTGACTCTGAATGTGGCCAGAGATTCCGGAGTTGGTGTCTGAAT--TCTCCCGCATGGC 4669  
 QY 1095 rGgLyArgPro 1098  
 Db 4670 CAGGACCCCC 4680

RESULT 14  
 US-09-854-356-9  
 ; Sequence 9, Application US/09854356  
 ; Patent No. US20020177567A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Corixa Corporation  
 ; APPLICANT: SmithKline Beecham Biologicals S. A.  
 ; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
 ; FILE REFERENCE: 014058-009810PC  
 ; CURRENT APPLICATION NUMBER: US/09/854,356  
 ; CURRENT FILING DATE: 2001-05-09  
 ; PRIOR APPLICATION NUMBER: US 09/493,480  
 ; PRIOR FILING DATE: 2000-01-28  
 ; PRIOR APPLICATION NUMBER: US 60/117,976  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 9  
 ; LENGTH: 3768  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(3768)  
 ; OTHER INFORMATION: human HER-2/neu protein  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(1959)  
 ; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (2026)..(3765)  
 ; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (2968)..(3765)  
 ; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (3144)  
 ; OTHER INFORMATION: preferred portion of the phosphorylation domain  
 ; OTHER INFORMATION: (delta PD) of human HER-2/neu  
 US-09-854-356-9

Alignment Scores:  
 Pred. No.: 1.16e-39 Length: 3768  
 Score: 474.00 Matches: 261  
 Percent Similarity: 36.85% Conservative: 116  
 Best Local Similarity: 25.51% Mismatches: 337  
 Query Match: 8.09% Indels: 312  
 DB: 9 Gaps: 46

US-09-397-967-16 (1-1099) x US-09-854-356-9 (1-3768)  
 QY 223 ProCysGlyArg-----LeuProGlyArgProTyr---AlaLeuMetAlaLysTyrLeu 240  
 Db 381 CCTGTCTACAGGGGCTCCAGGAGGCTCCGGAGGCTGACCTTCCAGACCTTCACAGA 440  
 QY 241 AspleuGluArg-----LeuHIsProAlaAlaThrTrpGluThrPheArgValGlyLeu 258  
 Db 441 GATCTGTAAAGAGGGGCTTGTATCCAGC----- 470  
 QY 259 ProGlyAlaGlnGluGlyProGlyLeuLeuArgValAlaGlyAspAsnGlyTleProTrp 278  
 Db 471 -----GAACCCCGACGCTCTGCTA-----CCAGA 494  
 QY 279 SerSerAsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspValSer 298  
 Db 495 CAC-----GATTTTGTGGA----- 509  
 QY 299 IleAsnGlnAlaProAlaArgValGlyProAlaGlyGlnHIsArgLeuValThrValThrArg 318  
 Db 510 --GGACATCTTCCACAGACACACACAGCTGGCTCTAC----- 545  
 QY 319 MetAspGlyHIs----- 322  
 Db 546 ACTGATAGACACACACACCGCTCTGGGCTGCCACCCCTTTCTCGATGTGTAAAGGCTC 605  
 QY 323 --IleLeuGlnAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuVal 341  
 Db 606 CCGCTGCTGGGAGAGAGATTCTGAGAGATTGTCACAGC--CTGAC----- 648  
 QY 342 AspGlyTyrPheArgLeuIleCysAspSerArgHIsTyrPheCysAlaValAlaPro 361  
 Db 649 -----CCGACTGCTGTGCCGCTGGCTGTGCCCTGCAGAGGCGCACTGCC 666  
 QY 362 ProArgLeuLeuGlnGlu----- 368  
 Db 697 ACTGACTGCTCCATGACAGCTGTGCTGCTGCCGCTGCCACGGGCCCAAGCACTGTGACTGC 756  
 QY 369 -----AlaAspValCys-----HIsGlyProIleThrLeu 378  
 Db 757 CTGGCTGCCCTCCACTTCAACACAGAGTGCATCTGTGACGTGCACATGCCACCGCTGTGTC 816  
 QY 379 AspPheAlaIleHIsLysLeuLysAlaAlaGlySerLeuProGlyTyrTleLeuArg 398  
 Db 817 ACCTACACACACACACACCTTTGAGTCCATGCCCAATCCCGAGGCGGCTAT----- 867  
 QY 399 ArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGly 418  
 Db 868 -----ACATTCGGCGCCAGCTGTGTGACTGCTGCTCCCTACAACTTACTTCT 915  
 QY 419 ProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGly 438  
 Db 916 ACGGACGTGGGATCTGTGACCTT--CGTGTGCCCT----- 950  
 QY 439 LeuSerGlnProHIsArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsn--SerGlyL 458  
 Db 951 ---GCACAAACCAAGAGTGACAGACAGGA-----TGAACACACACGGGTG 992  
 QY 458 euArg-----ValAspGlyAlaAlaLeuTyrLeuThrSerCysCys- 471  
 Db 993 TGAGAGTGCAGACAGCCCTGTGCCGAGTGTGCTGTGCTGGGCATGAGACCTTTCG 1052  
 QY 472 -----AlaProArgProLysGlyLysSerAsnLeuIleValAlaArgArg- 486  
 Db 1053 AGAGGTAGGGCAGTTACCACTGATGCCAATATCCAGAGAT--TTGCTGCTGCACAGAAAT 1109  
 QY 487 -----GlyCysAsnProAlaProAlaP 494  
 Db 1110 CTTTGGAGCCTGGCATTTCTGCGGAGAGCTTTGATGGGAGCCAGCCTCCAAACATGTC 1169  
 QY 494 roGlyCysSerProSer-----CysCysAlaLeu-----ThrGlnLeuSerP 508

```

QY      800  aclyaaclaeuylrualacysgslmsprrolaallernegluclualgnlsleuylsty  820
           |||||:::  |||  ::::||  ||:::  ||:::  ||:::
Db      2128  ---CGCGAGATG-----CGAGATCTGAAGAGACGAGCTAGAGAA  2165
QY      820  rllserleueuglylsglyasrphesylservalagluileucysargtuyasrprole  840
           :::::|||||  |||  |||||:::||||  :::::  |||  :::::
Db      2166  GGTGAAGGTGCTTGGATCTGGCGCTTTTGGCAAGCTCACAAAGGCGATCGGATCCCTGA  2235
QY      840  ucllyaspsnthrlglyproleuvalalalygslne--glnlsservalproas  859
           ||:::||||  |||||:::||||  |||  :::::  |||  :::::
Db      2226  TGGGAGAAATGTGAATAATCCAGTGGCGCATCAAAAGTGTGAGGAAACACATCCSSCAA  2285
QY      859  rclnlnlrgrsphneglnatrguillieglileuylsrlalaleuhslerasrphell  879
           :::::  :::::  |||  :::::  :::::  |||  :::::
Db      2286  ACCCAACAACAAATACTTAGACAAACATACAGTAGGCGTGGTGGCGCTCCCATATGT  2345
QY      879  eualylstytrgrgllyualsertyrglrprogluatrglnsertleatrgluvalamegl  899
           :  ::  ||:::  :::::  :::::  :::::  :::::  :::::
Db      2346  CTCGGCGCTTGTGGGCATCTGCCG-----ACATCCAGGTGCACCTGTGACACA  2396
QY      899  utlyleuproserglcysleuatrgasrleuclnlatrgnlisatrgly--leuilsth  918
           :  ::::  |||||::|  ||  :::::  :::::  |||||  |||  :
Db      2397  GCTTATGCCCTATGATGGCGCTGCTTAGACACATGTCCGGGAAACCCGGCGCGCTGGCTG  2456
QY      918  rlsprargleuileuehealatrgrnllcylslysglmeteglutlytleuylalalatr  938
           :  |||||  ::::  |||||  |||||  |||||  |||||  |||||
Db      2457  CCAGCACTGCTGACAGTGGTGTATGACAGATGCCAAGGAGATGAGCTCTCGAGAGATGT  2516
QY      938  gtrgcyvalhlsatrgasrpeuallaalatrgasnlleleuvalglusergylualnlhsva  958
           :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      2517  GCGGCTGTACACAGGAGCTTGGCGCGCTGGCAACGTGCGTCAAGATCCCAACCAATGT  2576
QY      958  llysllealaspheglyleualalalyseuleuoproleuyllylsasrtyttrgylvala  978
           |||||  |||||:::||||  :::::  |||  ::::  |||
Db      2577  CAAATTTACAGACTTCGCGCTGGCTGGCTGGACATTTGACAGACAGATACCATGTC  2636
QY      978  lartgluproglinserrproilernhetrryrlarproglinsertleaserasrpnll  998
           :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      2637  A---GATGGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTGCATTTCTCCCGGCGC  2693
QY      998  epheserarglnsersrpaltrtserphegllyalvalleuylrgluileuherthty  1018
           ||:::  |||||:::||||  :::::  |||||  :::::  |||||  |||
Db      2694  GTTCCACCCACAGAGTGAATGATGATGATTTATGTTGTGTGATGTGTGGAGATGATGATTT  2753
QY      1018  rcyasrplyssercysserproseralaglupheleuatrgmetmetglrprogluatrgl  1038
           :  ::  |||  ::::  ::::  ::::  ::::  ::::  ::::
Db      2754  TGGGGCCAAA-----CCTTAGCA  2771
QY      1038  uglrproleuclysargleucluleuylleualagluclyatrgargleuoprotr  1058
           |||||  |||||:::  ::::  |||||  ::::  |||||  |||||  ||
Db      2772  TGGATGCCACCGCGGAGATCCCTGACCTGCTGGAAAAGGGAGACGGCTCCSSACCC  2831
QY      1058  cprothrcysbprothrgluvalaglnleuemetglnleucystrralarpogluarproh  1078
           |||  |||  ::::||  ::::||  |||||  |||||  ::::
Db      2832  CCCCATTCGCACCAATGTATGATGTACATGATGATGCAATATTTGGAGATGTATGATGTA  2891
QY      1078  saspargrprolahealatrheulserproglinsleuarprrogluatrgluyatrgt  1098
           |||||  |||  |||  ::::  |||  |||  |||  |||  |||
Db      2892  ATGTGGGCCAAGATTTCCGGGAGTTGTGTCTGAAT--TCTCCCGCATGCCACGGGACCCC  2949
QY      1098  o  1098
           |
Db      2950  C  2950

RESULT 15
US-09-930-125-1
; Sequence 1, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.

```

RESULT 15  
US-09-930-125-1  
; Sequence 1, Application US/09930125  
; Publication No. US20020193329A1  
; GENERAL INFORMATION:  
; APPLICANT: Hand-Zimmerman, Susan  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Foy, Teresa M.

APPLICANT: Lodes, Michael J.  
 APPLICANT: Kalos, Michael D.  
 APPLICANT: McNeill, Patricia D.  
 APPLICANT: Vedvick, Thomas S.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
 FILE REFERENCE: OF HER-2/NEU-ASSOCIATED MALIGNANCIES  
 CURRENT APPLICATION NUMBER: 210121.344  
 CURRENT FILING DATE: US/09/930.125  
 NUMBER OF SEQ ID NOS: 25  
 SOFTWARE: FASTSEQ for Windows Version 3.0  
 SEQ ID NO 1  
 LENGTH: 3768  
 TYPE: DNA  
 ORGANISM: Homo sapien  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)...(3765)  
 US-09-930-125-1

## Alignment Scores:

Pred. No.: 1.16e-39 Length: 3768  
 Score: 474.00 Matches: 261  
 Percent Similarity: 36.85% Conservative: 116  
 Best Local Similarity: 25.51% Mismatches: 337  
 Query Match: 8.09% Indels: 312  
 DB: 9 Gaps: 46

US-09-397-967-16 (1-1099) x US-09-930-125-1 (1-3768)

QY 223 ProCysGlyArg---LeuProGlyArgProTy---AlaLeuMetAlaLysTyrIleLeu 240  
 DB 381 CCTGTACAGAGGGCTCCAGAGAGCTGGGAGCGAGCTTCAACCCCTCAGAGA 440  
 QY 241 AspLeuGluArg---LeuHisProAlaIaIaThrGluThrPheArgValGlyLeu 258  
 DB 441 GATCTTGAAAGAGGGGCTTGATCCAGCG----- 470  
 QY 259 ProGlyAlaGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyr 278  
 DB 471 -----GAMCCCCCACTCTGCTA-----CCAGA 494  
 QY 279 SerSerAsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspValSer 298  
 DB 495 CAC-----GATTTTGTGGAA----- 509  
 QY 299 IleAsnGlnAlaProArgValGlyProAlaGlyGlnHisArgLeuValThrValThrArg 318  
 DB 510 ---GGACATCTTCCACAGAACACACAGCTGGCTTCAC----- 545  
 QY 319 MetAspGlyHis----- 322  
 DB 546 ACTGATACACCAACCGCTCTCGGCTGCACACCCCTGTTCTCCAGTGTGAAGGCTTC 605  
 QY 323 ---IleLeuGlnAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuVal 341  
 DB 606 CCGCTGGTGGGAGAGAGCTTCTAGAGATTCAGAG--CTGACG----- 648  
 QY 342 AspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaPro 361  
 DB 649 -----CGCACTGTCTGTGCGGCTGTGGCTGCCCGCTGCACAAAGGGCCACTGGCCC 696  
 QY 362 ProArgLeuLeuGluGluGlu----- 368  
 DB 697 ACTGACGTGCGCATGACAGCTGTCTGCTGCCGCTGCACGGGCCCAAGCAGCTGTGACTGC 756  
 QY 369 -----AlaAspValCys-----HisGlyProIleThrIleu 378  
 DB 757 CTGGCCCTGCCCTTCAACACACAGTGGCATCTGTGACTGTGACTGCCCAAGCCCTGCTC 816  
 QY 379 AspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArg 398  
 DB 817 ACCTACACACAGACAGCTTTGAGTGCATGCCAATCCCGAGGCGCGGTAT----- 867

QY 399 ArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGly 418  
 DB 868 -----ACATGGGGCCAGCTGTGTGAGCTGCTCTCCCTCACTACTTCTTCT 915  
 QY 419 ProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGly 438  
 DB 916 ACGAGCGTGGAGTCTCTGACCT--CGTCTGCCCT----- 950  
 QY 439 LeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsn--SerGlyL 458  
 DB 951 ---GCACAACCAAGAGTGCACAGAGA-----TGGAACACAGCGGTC 992  
 QY 458 euArg-----ValAspGlyAlaAlaLeuTyrLeuThrSerCysCys- 471  
 DB 993 TGAGAGTGCAGCAAGCCCTGTGCCCGAGTGTCTGATGCTGTGGCATGAGCAGCATTTGGG 1052  
 QY 472 -----AlaProArgProLysGluLysSerAsnLeuIleValAlaArg- 486  
 DB 1053 AGAGTGAGGGCAGTATTCACAGTGCATATTCACAGACT--TTGCTGGCTGCAGAAAGAT 1109  
 QY 487 -----GlyCysAsnProAlaProAlaP 494  
 DB 1110 CTTTGGAGCCTGGCATTTCTGCCGAGAGACTTTGATGGGAGCCGCTCCAGACACTGC 1169  
 QY 494 roGlyCysSerProSer-----CysCysAlaLeu-----ThrGlnLeuSer 508  
 DB 1170 CCGGCTCCAGCCAGAGCAGCTCCCACTGTTTGAAGACTTGGAAAGATCAGAGTTTACT 1229  
 QY 508 he-His-----ThrIlePro----- 512  
 DB 1230 ATACATCTCAGATGGCCGAGACGCTGCTGACTCAGCGCTTCAGAACCTGCAGACT 1289  
 QY 513 -----ThrAspSerLeuGluThr-----His 519  
 DB 1290 AATCCGGGAGCAGATTTCTGCACATGGCGCTACTGCTGACCTGCAAGAGGCTGGGAT 1349  
 QY 520 GluAsn-LeuGlnHisGlySerPheThrLysIlePheArgLysArgArgGluValVa 539  
 DB 1350 CAGCTGGCTGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1403  
 QY 539 LaspGlyLysThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArg 559  
 DB 1404 CCACCATACACACCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1463  
 QY 559 nCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHis 579  
 DB 1464 C-----CCGACCA 1472  
 QY 579 uValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnIlePheValTy 599  
 DB 1473 AGCTGTGCTGCACACCTGCACACCGCCGAGAGGAC-----GAGTGTGTGG 1517  
 QY 599 rLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerLysSerPro 619  
 DB 1518 CAGAGGCTGGCTGCCACACAGCTGTGCTGCCCGAGGAC-----TGTCTGGGG 1565  
 QY 619 sleuGlnValThrLys-----GlnLeuAlaTyrAlaLeuAsnTyrLeuGln 634  
 DB 1566 TCCAGGGCCCAACCCAGTGTCACTGACAGCCAGTTCTTCCGGGCGCAGAGTGGCTGGA 1625  
 QY 634 uAsp-----LysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeu 650  
 DB 1626 GGAATGCGAGTACTGACAGGGGCTCCCAAGGAGGTATGTGATGCCAGGACACTGTTTGC 1685  
 QY 650 uAlaArg-----GluGlyLysAsp----- 656  
 DB 1686 GTGGACACCTGATGTGCAGCCCGCCAGATGACCTGTTTGGACCGAGGCTGA 1745  
 QY 657 -----GlyAsnProProPheIleLysLeuSerAspPro-- 667  
 DB 1746 CCAATGTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1805

QY 668 -GlyValSerProThrValIleuSerLeuGluMetLeuThrAspArgIleProTyrValAl 687  
 DB 1806 CGGTGGAAACCTGACCTCTCTACATGCCATC-----TCGAAGTT 1847  
 QY 687 apProglu-----CysLeuGlnGluAlaGlnThrLeuCysLeuG1 700  
 DB 1848 TCCAGATGAGGAGGGCGATGCCAGCTTCCCATCACTGACCCACTCTCTGTGTGA 1907  
 QY 700 uAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaH1 720  
 DB 1908 CCTGGATGACAGAGGCTGCCGCC-----GAGCAGAGAGCCAGCC 1949  
 QY 720 sIleThrSerLeuGluProAlaLysLysLeuLysPheTrpGluAspGlnGlyGlnLeuPr 740  
 DB 1950 TCTGACGTCATCATCTCTCCGGTG-----GTTGGCATCTCTGT 1988  
 QY 740 oAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProG1 760  
 DB 1989 GGTGCTGCTTGGGGTGGTGGTCTTGGGATCTCATCAAG---CGACGGCAGCAGAGAT 2045  
 QY 760 YArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTy 780  
 DB 2046 CCGAGATGACAGATGCGGAGACTGCTGCAGAAACG----- 2082  
 QY 780 rGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAl 800  
 DB 2083 -GAGCTGTGTGAGCGCTGACACCTACGCGAGCATGCCACACAG----- 2127  
 QY 800 aGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuLysTyr 820  
 DB 2128 ----GCCAGATG-----CGATCTCGAAGAGAGAGCGCTGAGGAA 2165  
 QY 820 rIleSerLeuGlnGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLe 840  
 DB 2166 GGTGAAGTGTCTGTGATCTGGCGCTTGGCAGATCTACAAAGGCATCTGGATCCTGA 2225  
 QY 840 uGlyAspAsnThrGlyProLeuValAlaValLysGlnLeu---GlnHisSerValProAs 859  
 DB 2226 TGGGGAGATGTGAAATTCAGTGGCCATCAAGTGTGAGGAGAAACACATCCCCAA 2285  
 QY 859 pGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheI1 879  
 DB 2286 AGCCAAACAAGAAATCTTGACGAAAGCATGATGCTGTGGCTCCCATATGT 2345  
 QY 879 eValLysTyrArgLysValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetG1 899  
 DB 2346 CTCCCGCTTCTGGGCATCTGCCG-----ACATCCAGGTGACAGCTGTGACACA 2396  
 QY 899 uTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly--LeuHisTh 918  
 DB 2397 GCTTATGCTTATGCTGCTTACCATGTCGCGAAACCGCGAGCGCTGGGCTC 2456  
 QY 918 rAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaAr 938  
 DB 2457 CCAGGACTCTGTAAGTGTGTATGCAAGTGGCAAGGGATGAGTACTGTGAGAGATCT 2516  
 QY 938 gArgCysValHisArgAspLeuAlaIaArgAsnIleLeuValGluSerGluAlaHisVa 958  
 DB 2517 GCGGCTGTACACAGGAGCTTGGCCGCTCGGAACGTGCTGTCAAGAGTCCCAACCATCT 2576  
 QY 958 lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVa 978  
 DB 2577 CAAATTAACAGACTTGGGGGCTGCGCTGAGCATTGACGAGACAGATGACATGC 2636  
 QY 978 lArgGluProGlyLysIleProIlePheTrpTyrAlaProGlyLysLeuSerAspAsnI1 998  
 DB 2637 A---GATGGGGGCAAGGTGCCATCAAGTGAATGCGCTGAGTCCATCTCCGCGGCG 2693  
 QY 998 ePheSerArgGlnSerAspValTrpSerPheGlyValValleuTyrGluLeuPheThrTy 1018  
 DB 2694 GTTCACCCACAGATGATGTGTGATGATGTGATGCTGTGGAGCTGATGACTT 2753  
 QY 1018 rCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgG1 1038

DB 2754 TGGGGCCAAA-----CCTACGA 2771  
 QY 1038 uGlyProPheLeuCysArgLeuLeuGlnLeuAlaGlnGlyArgGluProProP1 1058  
 DB 2772 TGGGATCCCAAGCCCGGAGATCCCTGACCTGTGAAAAGGGGAGCGGCTGCCAACCC 2831  
 QY 1058 oProThrCysProThrGluValGlnGluMetGlnLeuCysTrpAlaProGluProH1 1078  
 DB 2832 CCCCATCTGCACCATGTGATGTCTACATGATCATGTGTCAATGTGTGATGATGACTCTGA 2891  
 QY 1078 sAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgPr 1098  
 DB 2892 ATGTGGCCAAAGATTCGGGAGTGTGTCTGAT--TCTCCGATGCGCAGGAGACCC 2949  
 QY 1098 o 1098  
 DB 2950 C 2950  
 Search completed: April 28, 2003, 18:58:40  
 Job time : 419.323 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 28, 2003, 13:51:29 ; Search time 150.669 Seconds

(without alignments)  
2236.936 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 5860

Sequence: 1 MAPSEETPLIPQRCSSLS.....RPAPATLSPLQDPLMRGPG 1099

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 segs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n\_model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool/US09397967/runat\_26042003\_182313\_10750/app\_query.fasta.1.1486  
-DB=Issued\_Patents\_NA -OEMT=fastcap -SUFFIX=rnt -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=unna40.cdl  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967 -ECGN\_1\_1=59 -runat\_26042003\_182313\_10750 -NCPU=6 -ICPU=3  
-NO\_XIPXY -NO\_MMAR -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued\_Patents\_NA:  
1: /cgn2\_6/ptodata/2/1na/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/1na/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/1na/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/1na/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/1na/PCUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/1na/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5223	89.1	4016	5	PCT-US95-08354A-1
2	4418	75.4	3807	1	US-08-357-598-1
3	4418	75.4	3807	2	US-09-003-289-1
4	4418	75.4	3807	5	PCT-US95-16435-1
5	2629	44.9	3435	4	US-09-046-158A-21
6	2621	44.7	4482	3	US-08-567-508C-1
7	2621	44.7	4482	3	US-09-196-480-1
8	2619.5	44.7	3629	3	US-08-097-997A-8
9	2619.5	44.7	3629	3	US-08-665-574C-8
10	2619.5	44.7	3629	3	US-08-946-894-8
11	2360	40.3	3495	1	US-08-446-038B-2
12	2360	40.3	3495	1	US-08-446-010B-2

13	2360	40.3	3495	1	US-08-805-445-2
14	2360	40.3	3495	2	US-08-064-067D-2
15	2360	40.3	3495	2	US-09-066-208-2
16	2360	40.3	3495	4	US-08-980-080-3
17	1904	32.5	3429	1	US-08-097-997A-10
18	1904	32.5	3429	3	US-08-665-574C-10
19	1904	32.5	3429	4	US-08-946-994-10
20	1904	32.5	4234	1	US-08-446-038B-1
21	1904	32.5	4234	1	US-08-446-010B-1
22	1904	32.5	4234	1	US-08-805-445-1
23	1904	32.5	4234	2	US-08-064-067D-1
24	1904	32.5	4234	2	US-09-066-208-1
25	1815.5	31.0	3561	1	US-08-097-997A-12
26	1815.5	31.0	3561	3	US-08-665-574C-12
27	1815.5	31.0	3561	4	US-08-946-994-12
28	482	8.2	3845	2	US-08-220-240A-4
29	475	8.1	4530	1	US-08-229-515A-9
30	475	8.1	4530	1	US-08-645-865-9
31	475	8.1	4530	1	US-09-167-322-4
32	474	8.1	3768	2	US-08-625-101-1
33	474	8.1	3768	2	US-08-356-786-1
34	474	8.1	4473	2	US-09-048-804-1
35	474	8.1	4473	4	US-09-056-105-26
36	466.5	8.0	5532	2	US-08-475-035-3
37	466.5	8.0	5532	4	US-09-676-610B-17
38	465.5	7.9	5484	4	US-09-632-580A-3
39	465.5	7.9	5501	1	US-08-484-438-1
40	465.5	7.9	5555	1	US-08-484-438-3
41	465	7.9	8083	4	US-09-383-630-4
42	465	7.9	8083	4	US-09-383-630-5
43	464.5	7.9	3955	1	US-08-229-515A-14
44	464.5	7.9	3955	1	US-08-645-865-14
45	462	7.9	4508	5	PCT-US93-06251-34

#### ALIGNMENTS

RESULT 1  
PCT-US95-08354A-1  
Sequence 1, Application PC/TUS9508354A  
GENERAL INFORMATION:  
APPLICANT: Temple University - Of The  
APPLICANT: Commonwealth System of Higher Education  
TITLE OF INVENTION: JAK3 PROTEIN TYROSINE  
TITLE OF INVENTION: KINASE AND DNA ENCODING THE SAME  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seidel, Gonda, Lavoigna  
ADDRESSEE: 6 Monaco, P.C.  
STREET: Suite 1800, Two Penn Center  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 kb  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/08354A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/272,368  
FILING DATE: 8 July 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480  
REFERENCE/DOCKET NUMBER: 6056-203 PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549

; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4016 nucleotides  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single stranded  
 ; TOPOLOGY: linear  
 ; PCT-US95-08354A-1

Alignment Scores:  
 Pred. No.: 0 Length: 4016  
 Score: 5223.00 Matches: 1060  
 Percent Similarity: 80.59% Conservative: 3  
 Best Local Similarity: 80.36% Mismatches: 32  
 Query Match: 89.13% Indels: 235  
 DB: Gaps: 6

US-09-397-967-16 (1-1099) x PCT-US95-08354A-1 (1-4016)

QY 1 MetAlaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
 DB 7 ATGGCACCCTCCAAAGTGAAGACACCTGTGATCCTCAGCGCTCTTACAGCTTCATCC 66  
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40  
 DB 67 TCACAGGAGAGAGCCCTGCATGTCTCTTCCCTCCGCGGAGCTGGGCTCCACAGCGA 126  
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuValArgAlaAlaGlyAla 60  
 DB 127 TTGCATCTCTCTTTGGGACTACTGTGAGATTATGTGGCAGCTGCCAAGGCC 186  
 QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
 DB 187 TGTGCGATCCGCTGCTTTATCATCTCGCTTTCGCTCGGCGACTGAGACTTCTTGC 246  
 QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
 DB 247 TGGTTTCCCAAGACCAATCTTCTGATAGAGACGTGACACTCAAGTCTTGCTTAC 306  
 QY 101 ArgLeuArgPheTyrPheProAspThrPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
 DB 307 AGGCTACGCTTTATTTCTCTGACTGCTGTTGGCTGGAGACATGTCACCGCTTGGGCTG 366  
 QY 121 ArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
 DB 367 CGCAAGATTTGACACAGTCGATCTTACTTACATGTTTAGAACATCTTTGCTCAG 426  
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLeuGluGlnGly 160  
 DB 427 CACCCAGTGAAGCTGTGAGTGGGCTCCCGGTGGGCTTACGATCAAGAGAGAGGGA 486  
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArg 180  
 DB 487 GAGTTCCTGAGCGCTGGCTGTGACTTGGCCAGATGGCTCGAGACAGCCAGCGC 546  
 QY 181 ProGlyGluLeuLeuLysThrValSerTyrIleValCysLeuProProSerLeuArgAsp 200  
 DB 547 CCAGAGAGAGCTGCTGAACAGCGCTCAGTACAAAGCCTGTCTCCGCCAGCTGGCCAT 606  
 QY 201 ValIleGlnGlyGlnAspPheValThrArgArgArgIleArgArgThrValValLeuAla 220  
 DB 607 GTGATCCAGGCGCAAACTTCTGTACACGACGACGCGCATCCGAGAGACGCTGTGGCG 666  
 QY 221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240  
 DB 667 CTGCG--GGCTGTGGTGGCTGCGCAGCGCGACGCTACAGGCTCATGGCAAGATAT-CTG 723  
 QY 241 AspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGly 260  
 DB 724 GACGTGAGACGGCTACATCCAGCGCCACACGAGACCTTCCGTGGGGCTCCCGGGC 783  
 QY 261 AlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProThrSerSer 280  
 DB 784 GCCCAGAGAGAGCCGGGCTTCTGCTGTGGGGGAGACAGCGCATCTCTCTGGAGCTCC 843

QY 281 AsnAsp----- 282  
 DB 844 GGGACCAGAGAGTGTGGGGCTGGGCTTGAGAACGGGGGTCCGGGGAGCGTGGCGG 903  
 QY 282 ----- 282  
 DB 904 TTAGTAAGAGAGCCGTCCGAACAATCAGAAAGTGCTAGTCCGGGGCGGCTGATTGC 963  
 QY 282 ----- 282  
 DB 964 ACTACGGGGCAGGGGGGGGCGGTCAATTTAAAGTGGTCCGGGATGAGAGCTGCCCGAC 1023  
 QY 282 ----- 282  
 DB 1024 GGAAGTACCTGGGGCGTGAAGTGAAGAGGTAGAGTATAGGGGGCGGGAGCAGAGAA 1083  
 QY 282 ----- 282  
 DB 1084 GGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1143  
 QY 282 ----- 282  
 DB 1144 ACCTGTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1203  
 QY 282 ----- 282  
 DB 1204 GGGTCAAGTCTGTAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1263  
 QY 282 ----- 282  
 DB 1264 TTTGCGAGAGAGGTTTCTGCTATAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1323  
 QY 282 ----- 282  
 DB 1324 GGCTGGCTGAGACTGCAAGATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1383  
 QY 283 ----- 283  
 DB 1384 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1443  
 QY 290 AspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGly 309  
 DB 1444 GACTTTCGGAATACGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1501  
 QY 310 GluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPhePro 329  
 DB 1502 AG-CACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
 QY 330 GlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCys 349  
 DB 1561 GGGCTGCTGAGGGCGCTGTCTTTTGTGGCCCTCGTGGATGGTACTTCGCCCTGATCTGC 1620  
 QY 350 AspSerArgHisTyrPheCysIleGluValAlaProProArgLeuLeuGluGluAla 369  
 DB 1621 GACTCCAGGATTTATTTCTGCAAGAGAGTGGCGCCAGCGCTGTGAGAGAGAGCGC 1680  
 QY 370 AspValCysHisGlyProIleThr----- 377  
 DB 1681 GAGCTGTGCTGATGAGCCCATCAGCTTAGAGAGCTAGTGGGTCAACAGTGGGATGGGGA 1740  
 QY 378 ----- 378  
 DB 1741 TGGCAGCGAGAGCTGTCTTGGTTGGACCTGGCTGTCTTGGCTAGTATGAGCTTT 1800  
 QY 381 AlaIleHisIleLysLeuValAlaGlySerLeuProGlyThrTyrIleLeuArgArgSer 400  
 DB 1801 GCCATCCCAAGATGAGAGCGCGCTGCGTCCCTCCAGGACCTATATCTCCCGCGCAGC 1860  
 QY 401 ProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420  
 DB 1861 CCGCAGGACTATGACAGCTTCTTCTTACCGCTGCGCTGAGACTCTCTTGGGCCCGGAC 1920





Tue Apr 29 06:04:28 2003

ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/357,558  
FILING DATE: 15-DEC-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/033001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3807 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
IS-08-357-598-1

Alignment Scores:	
Pred. No.:	Length:
Score: 0	4418.00
Percent Similarity: 86.9%	Matches: 870
Best Local Similarity: 81.16%	Conservative: 62
Query Match: 75.3%	Mismatches: 130
DB: 1	Indels: 12
	Gaps: 6

Qy	1	MetlaPrProSerGIuInThProleuIleProIlnArgSerCysSerIleuSer	20
Db	168	ATGGACCTCCAAATGAAGAGACCCCTATCCCTAGGGTCATCACCTCTTCTCC	227
Qy	21	SerGIuInlaGIaIAlaLeuHisValIleuLeuProGArgIlyProGIyProIlnArg	40
Db	228	ACGGAGGCTGGTGGCCCTGCATGCTCTGCGCCCTGGGGGCCCCCGCCACGCC	287
Qy	41	LeuSerPheSerPheGIaSPtyrIleuAlaGIaIAspLeuCysValaArgaIalaIAspAla	60
Db	288	CTATCTTCTCTCTTGGGGACACACTGGCTGAGGACCTGTGCGTGCAGCTGCACAGGCC	347
Qy	61	CysGIyIleuProValIlyrHisSerIleuPheAlaIleuAlaThrGluAspPheSerCys	80
Db	348	AGCCGATCTCCCTGTGTACCACTCCCTCTTCTCTGTGGCAGGAGACCTGTCTGC	407
Qy	81	TrpPheProPheSerHisIlePheCysIleGluAspValAspThrGlnValIleuValIy	100
Db	408	TGGTT-CCCCGAGACCACATCTTCTCCGTGAGGATCCAGACCCCAACATCTCGTGA	466
Qy	100	ArgIleuAlaArgPheIlyrPheProAspIlyrPheGIyIleuGIuInThCysHisArgPheGIyIle	120
Db	467	CAGATTCTGCTTACTCTCCCAATTGGTTGGCTGGAGAGGACCAACGGCTTGGGGCT	526
Qy	120	uArgIAspIleuThrSerAlaIleIleuAspLeuHisValIleuGIuInThIleuPheAlaGI	140
Db	527	ACGCAAGATTTGGCCAGTGTATCTTGAACCTCCAGTCTCTGGAGCAACCTCTTGGCCA	586
Qy	140	HisArgSerAspLeuValSerGIyAlaArgLeuProValGIyLeuSerMetIysGIuInGI	160
Db	587	GCACGACAGTACCTGTGATGGGGCCCTCCCGCCGGGCTCAGTCTCAAGAGAGCAGG	646

us-09-397-967-16.rni

Page 4

QY	160	YGLHLEUENSERLEUAlAValleuAspLeuAlaIagMeAlaArgIuInlaInr	180
Db	647	TGAGTGTCCAGCCCTGGCCGTGTGGACCTGGCCCGATGGCGGAGACGCCACAGC	706
QY	180	gProGIuLeuLeuLysThrValSerTYrLYsAlaCysLeuProSerLeuArgAs	200
Db	707	GGGGAGAGAGCTGCTGAAGACTGTACACTACAGAGCCTGCTACCCCAAGCTGGCGCA	766
QY	200	pValIleInIuLyInAsnPhenAlaThrArgArgIleArgArgPhrAlaValLeuAl	220
Db	767	CTGTATCCAGAGGCTGAGACTTCGTAGCGGGAGGCGATTTCGGAGGCGGTGGAGAGCC	826
QY	220	AluLeuLeuProCysGLyArgLeuProGIArgPro-TYrAlaLeuMeAlaLYsTYrIle	240
Db	827	CCGTGG-CCGGGTGGCGGCTGCCAGGACAGCGGACATCGTCATAGGCCAATCATCA	885
QY	240	eUAspLeuGIuArgLeuHisProAlaAlaThrThrGIuThrPheArgValGIyLeuProG	260
Db	886	TGGACCTTGAGACGGCTGGATCCACCGGGGGCGGCGAGACTTCACATGGCGGCTCCGAG	945
QY	260	LYAla-----GInGIuInuProGIyLeuLeuArgValAlaGLyAspAsnGIYIleP	277
Db	946	GGGCGCTTGTTGGCCACGAGCGGCTGGGGGCTCTCCGCGTGGCTGGAGCGGGGATGG	100
QY	277	roTYrSerSerAsnAsp---GIuLeuPheGIuThrPheCysAspPheProGIuIleValA	296
Db	1006	CCTGGACCCACGAGAGACAGAGAGGGTCTCCACAGCCTTCTCGACTTCAGAAATGCTAG	106
QY	296	spValSerIleAsnGIuAlaProArgValGIyProAlaGIyGIuHisArgLeuValThy	316
Db	1066	ACATTAGCATAGACAGGCCCCGGCGCTTGGCCGGCGGAGAGACAGCGCTGGTCACTG	112
QY	316	AlThrArgMetAspLYHisIleLeuGIuAlaGIuPheProGIyLeuProGIuAlaLeuS	336
Db	1126	TYACAGCAGACAGACACACAGATTTTAAGGCGCGATTCCAGAGGCTGCCAGAGCTCTGG	118
QY	336	eRPheValAlaLeuValAspLYIyrPheArgLeuIleCysAspSerArgHisTYrPheC	356
Db	1186	CGTTGCTGGCCCTGTGGAGCGGTACTTCGGCTGACACAGAGACTCCAGACTCTCTCT	124
QY	356	YsLYsGIuValAlaProProArgLeuLeuGIuGIuAlaAspValCYHisGIyProI	376
Db	1246	GCAAGAGAGTG---GACCCGAGGCTGCTGGAGAGATGGCGGAGCATGTCACAGGCCCCA	130
QY	376	IeThrLeuAspPheAlaIleHisLYsLeuYsAlaAlaGIySerLeuProGIyThTYrI	396
Db	1303	TCACCTGTGACTTTGGCTCCATCCAGAGCTCAAGACTGGGGCTCAAGCTCCGCTCTATG	136
QY	396	IeLeuArgArgSerProGIuAspIyrAspSerPheLeuLeuThrAlaCysValGIuThP	416
Db	1363	TYTCTGGCGGATCCCCAGAGACTTGGACACTTCTCTCTCACTGTGTGTGCAGAAC	142
QY	416	roLeuGIyProAspTYrLYsGLyCysLeuIleArgGIuAspProSerArgIYAlaPheSerI	436
Db	1423	CCCTTGGCTCGATTATTAAGGGCTGCTCATCGCGGCACACCCACAGAGAACTTCTTTC	148
QY	436	eUValGIyLeuSerGIuProHisArgSerLeuArgGIuLeuAlaAlaCysTrpAsnS	456
Db	1483	TGGTTTGGCTTAGCGGACCCACAGCAAGTCTTCAGAGAGCTCTGGCAACTCGTGGGATG	154
QY	456	eRGIyLeuArgValAspGIyAlaAlaLeuTYrLeuThrSerCysCYsAlaProArgProI	476
Db	1543	GGGGGCTGACAGTGATGGGTGGCGAGACCCATCTTCCTGCTGATGCCAGACCCA	160
QY	476	YsGIuLYsSerAsnLeuIleValIAlaArgArgGIyCysAsnProAlaProAlaProGIyC	496
Db	1603	AAGAAATCTCAACCTGATTTGTGTCCAGAGAGGTCTACACACCCACCAATCATCTTGG	166
QY	496	YsSerPro---SerCysCYsAlaLeuThrGIuInuSerPheHisThIleProThAspS	515
Db	1663	TTCAAGCCCCATCCAAATACAGCTGAGTCAAGATGACATTTTCAAGATCCCTGCTGACA	172
QY	515	eTLeuGIuTYrHisGIuAsnLeuGIuHisGLySerPheThrLYsIlePheArgGIYserA	535

```

Db 1723 GCCGAGGATGATGAGAACTGGGCGCTTCACCAAGATTATACGGGGCTGTC 1782
Qy 535 rgaatgagualvalaspolyglutrhinaspsersgluvalleuylvalmetasp 555
Db 1783 GCCATGAGGTGGTGGATGGGAGGCCGGAAGACAGAGGTCTGCTGAGGTCAATGATG 1842
Qy 555 eAtarhNstArasncysmetgluSerPheleugluAalaserleuemetsergluA 575
Db 1843 CCAAGCAACAAGACTGATGATGATGATTCCTGGAAGCGCAGGTGATGAGCCAAAGTGT 1902
Qy 575 eTrrProhNstleuValleuNstglValCysMetAlaGlyAspserllemetValG 595
Db 1903 CGTACCGGCAATCTGCTGCTGCTCCACGGCTGTGCATGCGTGAAGACACAGCATGCTG 1962
Qy 595 lngluPheValtryleuGluValAlaIleasPmetryleuArglysArglyNstleuVal 615
Db 1963 AGGAATTTGTAACACTGGGGCCATAGACATGTATCTCGAANAAGTGGCCACTGGTGC 2022
Qy 615 eAlaserTrpLysleuGluValAlaThrLysGluLeuAlaValAlaAsnTryleuGlu 635
Db 2023 CAGCAGCTGGAAGCTGGAGGTGGTCAACAGCTGGCCCTACGCGCTCAACTATCTGAGG 2082
Qy 635 sryLysGlyLeuProhNstglLysNstAlaSerAlaArglyValleuLeuAlaArgGly 655
Db 2083 ACAAGGCTGTCCATGAGCAATGCTCTGCGCGAAGGTCTGCTGGCGGAGGGG 2142
Qy 655 LysArgLysProProPheIleLysLeuSerAspProGlyValSerProThrValLus 675
Db 2143 CTGATGGGAGGCCCTTCATCAAGCTGAGTGAACCTGGGGTCAAGCCCTGTGTTAA 2202
Qy 675 eTleuGluMetLeuThrAspArgLysProTrpValAlaProGluNstleuGluAlaG 695
Db 2203 GCGTGAAGATCTCACGACAGATCCCTGGTGGGCCCGGAGTGTCTCGGGAGGGCC 2262
Qy 695 lntPrleuCysleuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPhe 715
Db 2263 AGACACTTAAGCTTGAAGCTACAGAGTGGGCTTGGCGCCAGCGTCCGGAAGGTGTA 2322
Qy 715 lnaArgGlyProAlaHstIleThrSerleuGluProAlaLysLysLeuLysPheTrpGlu 735
Db 2323 GTGGCGTCACCATGCCCATGCTGCTCAGATCTGCTGTAAGAACTCAATTTTGTGAG 2382
Qy 735 spgIngluGluLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCys 755
Db 2383 ACCGCGACAGCTGTGCGGCCCAAGTGGACAGAGCTGGCCCTGTGATCAACAGTCA 2442
Qy 755 eTAlaLysAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuNstgl 775
Db 2443 TGGCTATGAGCGCGGTCCAGAGGCTCTCTTACAGGCGGTATGCTCAATTAAGTCT 2502
Qy 775 euIleThrSerAspTrpGluLeuLeuSerAspProThrProGlyIleProSerProAla 795
Db 2503 TCATCTCTCTGACTATGAGTCTCTCTCAGACACACAC---TGGTGGCTTGGCACTGCTG 2559
Qy 795 spgLeuCysValAlaGlyAlaGluLeuTrpAlaCysGluAspProAlaIlePheGlu 815
Db 2560 ATGGGCTGTG---AATGGTCCCAAGCTCTATGCTTCCCAAGACCCCAAGATCTTCCAG 2616
Qy 815 lnaArgHstleuLysTrpIleSerleuGluGlyLysGlyAsnPheGlySerValGluNst 835
Db 2617 AGACACACCTTAAGTACTCTCAGAGCTGGGCAAGGGCTTCTTGGCAGCGGAGGTGT 2676
Qy 835 ysaArgTrpAspProLeuGluLysAsnTrpGlyProLeuValAlaValLysGluLeuGln 855
Db 2677 GCCGCTATGAGCCGCTAGGCGACATACAGTGGCTTGGCTGCGCTGAACAGCTGCAAGC 2736
Qy 855 lsservalProAspGlnGluAlaArgAspPheGluAlaArgGluIleGlnIleLeuLysAla 875
Db 2737 ACACGCGGCGCAGACAGAGAGGACTTTCAGCGGAGATTCAGATCTCAAAAGCAGAC 2796
Qy 875 lsservalPheIleValLysTrpArgGlyValSerTrpGlyProGlyArgGlnSerleu 895

```

```

Db 2797 ACATGATTTATGTCAAGATATCTGTGTACSTATATGCGCGCGGCGGACAGCCCTG 2856
Qy 895 rgleuValMetGluTrpLeuProSerGlyCysLeuArgAspLeuGluNstArgHstArg 915
Db 2857 CGCTGTATGAGTACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2916
Qy 915 lyleuNstHstAspArgleuLeuPheAlaTrpGluIleCysLysGlyMetGluTrp 935
Db 2917 GCGTGAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2976
Qy 935 euGluAlaArgArgCysValNstAspArgleuAlaAlaArgAsnIleLeuValGluSer 955
Db 2977 TGGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3036
Qy 955 lualaHstValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLys 975
Db 3037 AGGCACAGCTCAAGTCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3096
Qy 975 yTrpValValArgGluProGlyGlnSerProIlePheTrpTrpAlaProGluSerLus 995
Db 3097 ACTAGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3156
Qy 995 eAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValleuTrpGlu 1015
Db 3157 CGGACACATCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3216
Qy 1015 euPheThrTrpCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1035
Db 3217 TCTTACCTACTAGGACAAACACTGACAGCCCTCGCGGAGTCTCGCGGAGTATGGAGT 3276
Qy 1035 roGluArgGluGlyProProleuCysArgleuLeuGluLeuAlaGluGlyArgGly 1055
Db 3277 GTGACCGGAGATGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3336
Qy 1055 euProProProProProProProProProProProProProProProProProPro 1065
Db 3337 TGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3368

```

## RESULT 3

US-09-003-289-1

Sequence 1, Application US/09003289

Patent No. 5916792

GENERAL INFORMATION:

APPLICANT: Civin, Curt I.

APPLICANT: Small, Donald

TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish &amp; Richardson P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/003, 289

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/357, 598

FILING DATE: 15-DEC-1994

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07265/033001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099



```

Qy 635 sPLySGlyLeuProHISglYAsnValSerAlaArgLySValLeuLeuAlaArgLUGlyG 655
Db 2083 ACAAGGCGCTGTCCATGGCAATGTCTGCCCCGGAAGGTGCTCGGCTCGGAGAGGG 2142
Qy 655 LAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuS 675
Db 2143 CTGATGGGAGCCCGCTTCATCAACGTGATGACCTGGGCTGACCCCGCTGTAA 2202
Qy 675 eRLeuGLueLLeuThrAspArgIleProTrpValAlaProGlyLysLeuGLInGlyAlaG 695
Db 2203 GCCTGGAGATGCTCACCGCAGAGATCCCTGGGTGGCCCCCGAGTGTCTCCGGAGGCC 2262
Qy 695 LInThrLeuGlySLeuGLuaAlaAspLySTrpGlyPheGlyAlaThrTrpGlyValPheG 715
Db 2263 ACAGACTTAGCTTGGAACTGACAAGTGGGGCTTGGGCCCAAGCGTCTGGGAAGTGTAA 2322
Qy 715 LInArgGlyProAlaHISIleThrSerLeuGLuProAlaLysLysLeuLysPheTyrgL 735
Db 2323 GTGGCGTCAACATGCCCATCAGTGGCTGATGCTGCTAAGAACTCCAAATTTATGAGG 2382
Qy 735 sPGLInGlyLInLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrInGlyS 755
Db 2383 ACCGGAGCAGCTGTGCGGCCCAAGTGGACAGAGCTGGCCCTGTATTCAGAGTGA 2442
Qy 755 eValaTyArgProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyL 775
Db 2443 TGGCCTATGAGCGCGGTGACAGAGCCCTCTTACGAGCCGTACATGCTGACCTCAATAG 2502
Qy 775 euIleThrSerAspTyrgLuleuLeuSerAspProThrProGlyIleProSerProAlaG 795
Db 2503 TCATCTCTTCAAGACTAGTGGCTCTCTGACACACAC--TGGTCCCTGGCACCTG 2559
Qy 795 sPGLuleuGlySValAlaGlyAlaGInLeuTyAlaCysGlnAspProAlaIlePheGL 815
Db 2560 ATGGGCTGG--AATGGTGGCCAGCTCTATGCCCTCCCAAGACCCAGATCTTCGAGG 2616
Qy 815 LInArgHISLeuLysTyrlIleSerLeuGlyLysGlyAsnPheGlySerValGlyLeu 835
Db 2617 AAGAGACCTCAAGTACATCTCAGAGCTGGGCAAGGGCTCTTGGCACCGTGGAGCTGT 2676
Qy 835 YAsArgTyArgProLeuGLYAsnAsnThrGlyProLeuValAlaValLysGInLeuGLIn 855
Db 2677 GCGGCTATACCCCGTAGCGCAGCATACAGTGGCTGTGGCCCTGTGAACAGCTGCAGC 2736
Qy 855 LAsSerValProAspGlnGlnArgAspPheGlnArgGlyIleGlnIleLeuLysAlaLeu 875
Db 2737 ACAGCGGGCCAGACAGGAGGAGCTTACCGGGAGATTCAGATCTCAAGACAGACAGC 2796
Qy 875 LAsSerAspPheIleValLysTyArgGlyValSerTyrgLProGlyArgGInSerLeu 895
Db 2797 ACAGTGAATTCATGTCAAGTATCGGTGATGAGCTATGGCCCGGCCCGCCAGACCCCTG 2856
Qy 895 rGLeuValMeGILuTyrgLeuProSerGlyCysLeuArgAspLeuGlnArgHISArgG 915
Db 2857 CGCTGTGATGAGTACCTGCGCCAGCGGTCTTCCGACACTTCTCGACAGCCGACCGG 2916
Qy 915 LuleuHISThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMeGILuTyrgL 935
Db 2917 GCCTGATGAGCAGCGGCTCTTCTATCTTCCAGATGTGCAAGGGCAGTGGAGTACC 2976
Qy 935 euGlyAlaArgArgGlyValHISArgAspPheAlaAlaArgAsnIleLeuValGlySerG 955
Db 2977 TGGGCTCCCGCGCTGCTGACCGCGACCTGGCGCCGGAACATCTCTGCGGAGAGCG 3036
Qy 955 LAlaHISValLysLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGLYLysAsp 975
Db 3037 AGGCACAGCTCAAGATCGCTGCTGCGGCTGAGTACCTGCGGCTTGCACAAAGACT 3096
Qy 975 YTrTyValValArgGluProGlyLInSerProIlePheTrpTyrgLAlaProGInSerLeuS 995
Db 3097 ACTAGTGTGTCGCGAGCGACGAGCCAGACCCCATTTCTGTGATGCCCGCGAATCCCTCT 3156

```

```

Qy 995 eRAspAsnIlePheSerArgInSerAspValTrpSerPheGlyValValLeuTyrgL 1015
Db 3157 CGGACACACTTCTCTGCGCCAGCTCAGAGCTGTGAGCTTGGGGTCTGCTGTACGAGC 3216
Qy 1015 euPheThrTyArgAspLySserCysSerProSerAlaGluPheLeuArgMeGILY 1035
Db 3217 TCTTCACTACAGCACAAGAGCTGAGCCCTCGGCGAGTTCCTGCGGATGATGGAT 3276
Qy 1035 roGLuArgGluLysProProLeuGlyArgLeuGLuleuLeuAlaGlyArgTyrgL 1055
Db 3277 GTGACGGGATGTCCCGCGCTCTGCGGCTTGTGAACCTGTGAGAGGAGGCGCAGAGC 3336
Qy 1055 euProProProProThrCysProThrGluVal 1065
Db 3337 TGCCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3368

RESULT 4
PCT-US95-16435-1
; Sequence 1, Application PC/TUS9516435
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16435
; FILING DATE: 15-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halle, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/033W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; PCT-US95-16435-1

Alignment Scores:
Pred. No.: 0 Length: 3807
Score: 4418.00 Matches: 870
Percent Similarity: 86.94% Conservative: 62
Best Local Similarity: 81.16% Mismatches: 130
Query Match: 75.39% Indels: 12
DB: 5 Gaps: 6

US-09-397-967-16 (1-1099) x PCT-US95-16435-1 (1-3807)
Qy 1 MetaAspProSerGlnGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 168 ATGGACCTCCAAAGTGAAGAGAGCGCCCTGATCCCTCAGCGTTCATGCGACGCTTTC 227
Qy 21 SerGluAlaGlyAlaLeuHISValLeuLeuProProAlaGlyGlyProGlyProGlyAla 40
Db 228 ACAGAGGCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 287

```







Oy	177	nalaglnatgProGlygluLeuleuLysThyValSerTyrLysAlaCysLeuProProse	197
Db	582	CGATGCAAAACCCCTGCGCATCTGTAACTTAACCTTAACGTCACAAAGACATTCTTACCAAAATG	641
Oy	197	rLeuArAspValLlLlGlnIglYlAnsphValThrrArgArgIleArgIthrVa	217
Db	642	TATTGAGCAAAATGCCAAGACTACATCTTTTGGACAAAGGAAGGATTAAGGTACAGATT	701
Oy	217	lValleAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaL	237
Db	702	TCCGAGATTATTATCCAAACATTCAGCCAAATGCAAAACCCATCCGAAACCTGGAAACTTAA	761
Oy	237	sTyrIleLeuAspLeuGlnArgLeuLisProAlaIleThrIthrGlnThrPheArgIgl	257
Db	762	GTAATCTTAAATCTGGAAACTGCACTGCTGCTTTCACAGAGAAATTTGAAGTAA	821
Oy	257	ylLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs	273
Db	822	AGAACTCGGAAGTGGTCCCTCAGGTGAGAGATTTTTCACACATTATTAATCTGGAAA	881
Oy	273	pasnGlyIleProTyrSerSer-----AsnAspGluLe	284
Db	882	CGGTGAATTCAGTGGTCAAGAGGAACATAAAGAAATGAGACTGACAGCAACAGGA	941
Oy	284	uPheGlnThrPheCysAspPheProGluLlLeValAspAlaSerLleasnGlnAlaProAr	304
Db	942	TTTACAGTATATATGGCATTTTCTTAATATTAATGATGCAATATTAAGCAACAAACCA	1001
Oy	304	gValGlyProAlaGlyGlnHisArgLeuValThrValThrArgMetAspGlyHisLleLe	324
Db	1002	AGAGGG---TCAAAATAAAGCCGAGTGTACTATCCATTCATAGCAAGATGGTAAAAATCT	1058
Oy	324	uGluAlaGluPheProGlyLeuPheProGluLlLeuSerPheValAlaLeuValAspGlyTy	344
Db	1059	GGAAATTGCACTTAGTCATTAAGGAGACTTTGCTTTCGATGTCATTAAATTTGATGGATA	1118
Oy	344	rPheArgLeuLlLeCysAspSerArgHisTyrPheCysLysGluValAlaIleProProAlaGle	364
Db	1119	TTATGATTAACTCAGATGATCAGATCATATTAACCTCTGTAAAGAGTAGGACCTCCAGCCGT	1178
Oy	364	uLeuGluGluGlnLlLeAspArgLysHisGlyProLleThrLleuAspPheAlaIleHisLys	384
Db	1179	CGTTGAAAATATACAAAGCAACTGTACTGCGCCAAATTTGCATGAGATTTGGCATATAGTAA	1238
Oy	384	sLeuLysAlaAlaGlySerLeuProGlyThrTyrlleLeuArgArgSerProGlnAspTy	404
Db	1239	ACTGGAAGAAACAGAGTAAATGAGATGAGCTGATGTACTTCATGATCAGCTCAAGAGACTT	1298
Oy	404	rAspSerPheLeuLeuThrLlAcCysValGlnThrProLeuGlnProAspTyrLysGlyCy	424
Db	1299	TAAATAAATATTTTTCGACTTTTGCTGTGCGAGCGAGAAATGTGCATTTGATTAATAACATCG	1358
Oy	424	sLeuLleArgLlnAspProSerArgLylalPheSerLeuValGlyLeuSerGlnProHisAr	444
Db	1359	TTTGATTACAAAATAATGAGATGAGAGTCAACCTCAGCGGACAAAGAAACACTTCAG	1418
Oy	444	gSerLeuArgGlnLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAlaL	464
Db	1419	CAGTCTTAAAGATCTTTTGAATTTGTTACCAAGATGAGAACTGTCCGTCCAGACAAATATAT	1478
Oy	464	AlaLeuTyrlleuThrSerCysCysAlaProArgProGlyGlyLysSerAsnLeuLlLeVala	484
Db	1479	TTTCCAGTTTACTAAATGCTTCCGCCAAAGCCAAAAGATTAATCAAACTTCTAGTCTT	1538
Oy	484	lArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAlaLe	503
Db	1539	CAGAACGAAATGGTCTTGTGTGATGCACCAACCTCAGCAACATTCACAGAGCCATCATAT	1598
Oy	503	uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGlnTrpHisGlnAsnLeuLl	523
Db	1599	GAAACAAATGGTGTTCACAAAATCAGAAATGAAGATTTGATATTTTAATGAAAAGCTTGG	1655

QY	523	YHIGLYSerPheThrLysLLePheArgGLysERArgRgGLuValValAsp---	GLIGL	542
Db	1659	CCAAAGCACTTTTACAAAGATTTTTTAAAGCGCTAGCAAGAAAGTAGGAGACTGCGGTCA		1718
QY	542	uThHisSerPheLysValLeuLeuLysValMetAspSerLArgHisArgAsnCysMetGI		562
		:::		
Db	1719	ACTGCATCAACACAGAACTCTTTTAAAGTTCGGTAAAGCAGACAGAAACTATTGAGA		1778
QY	562	uSerPheLeuLysLValAlaSerLeuMetSerGLuValSerTYrProHisLysValLeuLe		582
Db	1779	GTCCTTCTTTGCAACACACAGATGATGATGAGCAAGCTTCTGCACAGCACTTTGGTTTTAA		1838
QY	582	uHisGLysValCysMetLLeuLysVal---SerLeuMetValGlnGluPheValTYrLeuGL		601
Db	1839	TTATGAGATATGATGCTCTGCTGGAGACAGCAATATTCTGCTTCAGAGCTTGTAAAAATTGG		1898
QY	601	ValAlaLeuAspMetTYrLeuLysArgLysArgGLuHisLysValSerAlaSerTYrPheLysLeuGI		621
Db	1899	ATCATCATATCATATTCTGAAAAAGAAATAAAAATTGTAAATATTTATGGAACCTTGA		1958
QY	621	nValThrLysGlnLeuLysLArgValAlaLeuLysnTYrLeuGluAsnTYrLysGLuPheLysProHisGL		641
Db	1959	ACTTCTTAAACAGTTGGCATGGGCGCATGCTTTTCTTGAAGAAACACCTTATTATTCAG		2018
QY	641	YAsnValSerLysLLeuArgLysLeuLeuValAlaArgGLuGlyGly-----AspGLysAsnPr		659
Db	2019	GAATGATATGTCSCAAAATATATCTGCTTATACAGAGAAAGACAGAGAACAGGAAATCC		2078
QY	659	oProPheLLeuLysLeuSerAspProGlyValSerProThrLysLLeuSerLeuGluMetLe		679
Db	2079	TCCTTTTCATCAACACTGATGTGATCTGTCATATATTACAGTTTGGCCAAAGACATCT		2138
QY	679	uThrAspArgLLeuProThrValAlaProGluCysLeuGlnGluAlaGlnThrLysCysLe		699
		:::		
Db	2139	TCAGAGAGAAATACCATGCGGTACACCTGAAATCATTTAAATCCAAAAATTTAAATTT		2198
QY	699	uGluValAspLysTYrPheGlyPheGlyAlaThrTrpTrpGluValPheGlnArgGLyProAl		719
Db	2199	GGCAACAGACAAATGGATTTGGTATCCACTTTGTGGAAATCTGCAGTGCAGAGAGATAA		2258
QY	719	aHisLLeuThrSerLeuGlnLysProAlaLysLysLeuLysPheTYrGLuAspGlnGlyLe		739
Db	2259	ACCTCTAGTGTCTGATGATCTCAAGAAAGATCAATTTTAAAGATAGGCAATCACT		2318
QY	739	uProAlaLeuLysTYrTrpThrGluLeuAlaGlyLeuLLeuThrGlnCysMetAlaTYrAspPr		759
Db	2319	TCCTTGACCAAGATGGCGAGAAATTAGCAACCTTATTAATATTTGATGATTATGAAAC		2378
QY	759	oGlyArgArgProSerPheArgAlaLLeuAlaArgAspLysAsnGlyLeuLLeuThrSerAs		779
Db	2379	AGATTTCAGGCTTCTTTCACAGGCATCATACGAGATCTTAACAGTTTCTTATTCACGAG		2438
QY	779	pTYrGluLeuLeuSerAspProThrProGlyLLeuProSerProArg---AspGluLeuCY		798
Db	2439	TTATGAGACTTTTAAACAGAA---AATGACATCTTACCAAAATATGAGATAGTGGCCCTACG		2495
QY	798	sValAlaGlyLagLLeuTYrAlaCysGlnAspProAlaLLeuPheGlnGluArgHisLe		818
Db	2496	GTTTTTGTGGTCC-----TTTGAAGACCGGGAGTCCACACATTTTGAAGAGAGCATTT		2549
QY	818	uLYsTYrLLeuSerLeuGluGlyLysGlyAsnPheGlySerValGluLeuCysArgTYrAs		838
Db	2550	GAATTTTCTACAGCAACTTGGCAAGAGGTAAATTTTGGAGATGTGGAATATGCGCGGTATCA		2605
QY	838	pProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPr		858
Db	2610	CCCTTCACAGGACACACACTGGGGAGCTGGCTGCTAAAAAGCTTCCAGCATATGATCTGA		2666
QY	858	oAspGlnGlnArgAspPheGlnArgGluLLeuGlnLLeuLysAlaLeuHisSerAspPh		878
Db	2670	AGAGCACTTAAGAGACTTTGAAAGAGGAATTTGAATCTCTGAATCCCTACAGCATGACAA		2725
QY	878	eLLeuLYsTYrArgGlyValSerTYrGlyProGlyArgGlnSerLeuArgLeuValnle		898



```

|||||
Db 2730 CATTGTAAGTCAAGGGAGTGTGCTACAGTCTGCTCGCGCTAACTAAATTAATTAAT 2789
Oy 898 tGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuH 917
Db 2790 GGAATATTACCATATGAGAGTTTACGAGACTATCTTCAATAAACATAAAGACGGATAA 2849
Oy 917 sThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlnAl 937
Db 2850 TCACATAAACCTTTCGACATACACATCTGACATATGCAGAGGGATGAGACTATCTTGTC 2909
Oy 937 sArgArgCysValHisArgAspLeuAlaHisArgAsnIleLeuValGluSerGluAlaHis 957
Db 2910 AAAAAGGATATATCCACAGGAGATCTGGCAACGAGAAATATATTGTCGAGAACAGAGAAC 2969
Oy 957 sValLysIleAlaAspPheGlyLeuAlaHisLeuLeuProLeuGlyLysAspTyrTyrVa 977
Db 2970 ACTTAAATTTGAGATTTTGGGTTAACCAAGCTTTCGCCACAGACAAAGAAATACTATTA 3029
Oy 977 lValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGlySerLeuSerAspAs 997
Db 3030 ACTAAAGAACTCGTGAAGATCCCATATTTGTATGCTCCAGATCACTGACAGAGAG 3089
Oy 997 nIlePheSerArgGlnSerAspValTrpSerPheGlyValAlaLeuTyrGluLeuPheH 1017
Db 3090 CAGATTTCCTGCGCTCAGATGTTTGAGCTTGGAGTGTCTGTATGAACCTTTCAC 3149
Oy 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArg 1037
Db 3150 ATACATTGAGAGAGATAAAGTCCACACAGCGAATTTATGCTATGATTTGGCAATGACAA 3209
Oy 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuPr 1056
Db 3210 ACAAGCACAGATAGCTGCTTCATTTGATAGAACCTTTGAAGCAATATGAAGAAATTAC 3269
Oy 1056 oProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl 1076
Db 3270 AAGACCAAGTGAATGCCCGATGAGATCTATATGATCATGACAGAGATGCTGGAAACATTA 3329
Oy 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
Db 3330 TGTAAATCAACGCCCTCTTACGAGATCTACCTTCGAGTGAATCAATA 3381

```

## RESULT 6

```

US-08-567-508C-1
: Sequence 1, Application US/08567508C
: Patent No. 5914393
:
: GENERAL INFORMATION:
: APPLICANT: Coleman, Roger
: TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: US
: ZIP: 94304
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/567,508C
: FILING DATE: 05-DEC-1995
:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.

```

```

: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-00490S
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-845-4166
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4482 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: LIBRARY: Placentia
: CLONE: 179527
:
: US-08-567-508C-1
:
: Alignment Scores:
: Pred. No.: 6,59e-234 Length: 4482
: Score: 2621.00 Matches: 532
: Percent Similarity: 66.03% Conservative: 193
: Best Local Similarity: 48.45% Mismatches: 342
: Query Match: 44.73% Indels: 31
: Gaps: 14
:
: US-09-397-967-16 (1-1099) x US-08-567-508C-1 (1-4482)
:
: Oy 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGl 39
: Db 447 TCCAGTTCTTCAGGT-----GTATCTTTACCATTTCCCTTGGAAATCTGAGCGACA 497
:
: Oy 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGlnAspLeuCysValArgAlaAla 59
: Db 498 TTATCTGACCTTTTCATGTCGGGAGATGTGTGAGAGAAATCTGATATGCTGCTTCTTA 557
:
: Oy 59 sAlaCysGlyIleLeuProValTyrHisSerIlePheAlaLeuAlaThrGlnAspPheSe 79
: Db 558 ACCTTGCTGATACACACCTGTGTATCATAAATATGTTTGGTTAAATAGAGTGAAGAAAG 617
:
: Oy 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99
: Db 618 GATCGGTATCCACCACCAACCATGCTTCTCATATAGATGATGACAGGCAATATGTA 677
:
: Oy 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluTyrCysHisAr 117
: Db 678 CTACAGATAAAGATTTTCTTCTTCGTTGATGATGATGATGATGATGATGATGATGATG 737
:
: Oy 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisLe 137
: Db 738 GCATGGAATATCTCGAGGTGCTGAAGCTCTCTTCTTGATGACTTGTGATGCTTCTTACT 797
:
: Oy 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLys 157
: Db 798 CTTTCTCAGTGGCGCATGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 857
:
: Oy 157 sGluGlnGlyIlePheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgLysGlu 177
: Db 858 AACACAGAGAAAGTATGCTTGGATGATGATGATGATGATGATGATGATGATGATGATG 917
:
: Oy 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProPse 197
: Db 918 CGATCAAAACCCCTGCGCATCTATTAACCTTCATGACGTCACAGACATTTCTTCCACAA 977
:
: Oy 197 rLeuArgAspValIleGlnGlyGlnAspPheValHisArgArgGlyLeuArgArgThrVa 217
: Db 978 TATTTGAGCAAGATCCAGAGATATCATATTTTGCACAGAGAACCAATAGAGTACAGATT 1037
:
: Oy 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaL 237
: Db 1038 TCGCAGATTTATTTCAGCATTCAGCCAAATGCAGCAATGCAGCAAACTTGAAACTTAA 1097
:
: Oy 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGl 257

```

Db 1098 GTATTATATACTGGAACCTGCGCTGCTTCTACAGAGAAATTTGAAGTAA 1157  
OY 257 YLEUPROGLYALA-----GInGUlUProGLYLeuLeuArgYAlaAGLYAs 273  
Db 1158 AGAACCTGGAAGTGCCTTCAGGTGAGAGATTTTGCACCACTATTAATTAAGTGA 1217  
OY 273 PAANGLYIleProTrpSerSer-----AsnAspGluLe 284  
Db 1218 CGGTGGAATTCAGTGCACAGAGGAAACATTAAGAAAGTACACTGCACAGACAGGA 1277  
OY 284 uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304  
Db 1278 TTACAGATTATATTCGATTTTCTTATATTAATGATGATGATTAAGACAGCAACCA 1337  
OY 304 gValGlyProAlaGlyGluHisArgLeuValIthrValThrArgMetAspGlyHisIleLe 324  
Db 1338 AGAGGGT---TCAAAATGAAAGCGAGTGTACATATCCATAGCAACATGATAAAATCT 1394  
OY 324 uGUlAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy 344  
Db 1395 GGAATTTGAACCTTAGCTCATTAAGGGAAGCTTTCTTCTGTCATTAATGATGATA 1454  
OY 344 rPheArgLeuIleCysAspSerArgHisIlyrPheCysGlyGluValAlaProProArgLe 364  
Db 1455 TTATAGATTAACTGCAGATCACATCATTTACCTCTGTAAGAAAGTACACCTCCAGCGGT 1514  
OY 364 uLeuGluGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisIly 384  
Db 1515 GCTTGAATAATATACAAAGCAACTGTCATGGCCCAATTCGATGGATTTTGCACTTAAGTAA 1574  
OY 384 sLeuValAlaAlaGlySerLeuProGlyThrTyIleLeuAlaArgArSerProGlnAspTy 404  
Db 1575 ACAGAAACAAAGCAAGTAAATCAGATGACAGTGTATCTAGTACGATGACGCTCTAAGGACTT 1634  
OY 404 rAspSerPheLeuLeuThrAlaCysValGlnIthrProLeuGlyProAspTyIlyGlyCy 424  
Db 1635 TAATTAATATTTTGTGCTTTGCTGTCGAGCGAAGAAATGTCAATTAATAATAAACACTG 1694  
OY 424 sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr 444  
Db 1695 TTTGATTACAAAATATAGATGAGTAGAGTACAACTTCAAGTGGCAAGAAAGAACTTCAG 1754  
OY 444 gSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAl 464  
Db 1755 CAGCTTAAAGATCTTTGATTTGTACCAAGATGAAAGCTTGCCTCAGACATATTAAT 1814  
OY 464 aLeuTyIleuThrSerGlyCysAlaProArgProGlyGlnIlySerAsnLeuIleVal 484  
Db 1815 TTTCAGATTACTAAATGCTGTCGCCCAAGCCAAAGATTAATCAAACTTCTAGTCTT 1874  
OY 484 lArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerGlyCysAlaLe 503  
Db 1875 CAGAACCAATGCTTTCTGTGATGACCAACCTCACAACATTAAGAGGCTTACTATAT 1934  
OY 503 uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuG 523  
Db 1935 GAACCAATGCTTTCTCACAAAATCAGAAATGAAGATTTGATTAATGAAGAACTTGG 1994  
OY 523 yHisGlySerPheThrIlyIlePheArgGlySerArgArgGluValAlaValAsp---GlyG 542  
Db 1995 CCAAGGCACTTTACAAAGATTTTAAAGCGCTACGAGAGAAAGATAGAGAGCTACGGTCA 2054  
OY 542 uThrHisAspSerGluValLeuLeuValMetAspSerArgHisArgAsnCysMetG 562  
Db 2055 ACTGCATGAACACAGATCTTTTAAAGTCTGTGATTAACACACAGAGACTATTCAGA 2114  
OY 562 uSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyIleProHisLeuValLeuLe 582  
Db 2115 GCTTTCTTTTAAGCAGCAAGATATGATGAGCAAGCTTCTCACAAGCATTTGTTTAA 2174  
OY 582 uHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyIleG 601  
Db 2175 TTATGAGATATGTCTGTGAGACAGAAATATTTCTGTTGAGAGTTTGAATAATTTGG 2234  
OY 601 yAlaIleAspMetLeuThrLeuArgIlyAsnGlyHisLeuValSerAlaSerTrpIlyLeuG 621  
Db 2235 ATCAGTATATATATCTGTAAGAAAGATTAATAATTTGATTAATATATATATGAACTTGA 2294  
OY 621 nValIthrGlnLeuAlaIlyrAlaLeuAsnTyIleGluAspIlyGlyLeuProHisG 641  
Db 2295 AGTTGCTTAACAGTGTGGCATGGCATGATTTTCTAGAGAAACACCTTATTCATG 2354  
OY 641 yAsnValSerAlaArgIlyValLeuLeuAlaArgGluGlyGly-----AspIlyAsnPr 659  
Db 2355 GAATGTATGTGCCAAAATATTTCTGTTATCAGAGAAAGACAGAGACAGAAATCC 2414  
OY 659 oProPheIleLeuSerSerPProGlyValSerProThrValLeuSerLeuGluMetLe 679  
Db 2415 TCCCTTATATTAACCTTAGTATCTGTCATAGTATATACGTTTGGCAAGACATCTT 2474  
OY 679 uThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe 699  
Db 2475 TCAGAGAGAAATACCATGGGTACACCTGATCATGTAAGTGAATAATTTAAATTT 2534  
OY 699 uGUlAlaAspIlyTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgIleProAl 719  
Db 2535 GGCACAGCAAAAGGAGTTTGTGTACCTTGTGGAAATCGCAGTGGAGAGATTA 2594  
OY 719 aHisIleThrSerLeuGluProAlaIlyIlyLeuIlySerPheTyGluAspGlnGlyIle 739  
Db 2595 ACCCTAGAGTCTGTGATTCGTAAGAAAGCAATTTATGAGATGAGCATGCGT 2654  
OY 739 uProAlaLeuIlyTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaIlyAspTr 759  
Db 2655 TCCGTGCACCAAGAGGGGCAAAATTAGCAACCTTAATTAATGTGATTAATGAAAC 2714  
OY 759 oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779  
Db 2715 AGATTTCAGGCTTCTTTCAGAGCATCATCAGATCTTAACAGTTTGTACTCCAGA 2774  
OY 779 pTyIleGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspIleuCy 798  
Db 2775 TTATGACACTTTAAACAGA---AATGACATGTTACCAATATGAGATGAGTACGCTTGG 2831  
OY 798 sValAlaGlyAlaGlnLeuTyIlyAlaCysGlnAspProAlaIlePheGluGluArgHisIle 818  
Db 2832 GTTTTGTGGTGC---TTTGAAGACGGGAGCTTCAACAGTTTGAAGAGACATTT 2885  
OY 818 uIlyTyIleSerLeuLeuGlyIlyGlyAsnPheGlySerValGlnLeuCysArgTyAs 838  
Db 2886 GAAATTTCTACAGCAACTTGGCAAGGTATTTTGGAGAGTGGAGATGGCGGTATGA 2945  
OY 838 pProLeuGlyAspAsnThrGlyProLeuValAlaValIlyGlnLeuGlnHisSerValPr 858  
Db 2946 CCCCTCAGAGACAAACACTGGGGAGTGGCGCTGTAAGAAAGCTTCAGATATGACTGA 3005  
OY 858 oAspGlnGluArgAspPheGlnArgGluIleGlnIleLeuIlyAlaLeuHisSerAspPr 878  
Db 3006 AGAGCACTTAAGAGACTTTGAAGAGAAATTTGAATCTGTAATCCCTACAGCATGCAA 3065  
OY 878 eIleValIlyTyIlyArgIlyValSerTyIlyGlyProGlyArgGlnSerLeuArgLeuValMe 898  
Db 3066 CATTTGTAAGTACAAAGGAGAGTGTCTACAGTGTGCGGAGTATCTAAATTAATTAAT 3125  
OY 898 tGluTyIleuProSerGlyCysLeuAlaArgAspLeuGlnArgHisArg---GlyLeuH 917  
Db 3126 GGAATATTTACCAATATGAAATTTACAGACTATCTCAAAACATTAAGAACAGATAGA 3185  
OY 917 sThrAspArgLeuLeuLeuPheAlaIlePrGlnIleCysIlyGlyMetGluTyIleGlyAl 937  
Db 3186 TCACATTAACCTTTGAGATGACATCTCAGATATGGAAGGATAGAGATATCTTGGTAC 3245  
OY 937 aArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaH 957  
Db 3246 AAAAAGATATTCACACAGAGATCTGCGACAGAGAAATATATATTTGTGGAGAAACGAG 3305



```

Db 1338 AGAGGCT---TCAAATGAAGACCCGAGTTGTAATCATATCAATAGCAAGATGGTAAAAATCT 1394
Oy 324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy 344
Db 1395 GGAATTTAACTTACCTCATTAAGGAAGCTTTGCTTCTGTCATTAATTAATGATGATA 1454
Oy 344 rPheArgLeuIleCysAspSerArgHisTyPheCysLysGluValAlaProProArgLe 364
Db 1455 TTATGATTTAACTGACATGACATCATTCCTCTGTAAGAAAGTACACCTCCAGCCGCT 1514
Oy 364 uLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLys 384
Db 1515 GCTTAAAAATATACAAACCACTGTCATGCCCCAATTCGATGGATTTTGGCATTAGTAA 1574
Oy 384 sLeuLysAlaAlaGlySerLeuProGlyThrTyIleLeuArgArgSerProGlnAspTy 404
Db 1575 ACTGAAGAAGACGAGTAATACAGACTGACCTGATGTACTGATCGATCCCTAAAGACTT 1634
Oy 404 rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyTyLysGlyCy 424
Db 1635 TAAATAAATATTTTGGACTTTTGTCTGCGAGCGAGAAATGTCATTGATTAATTAACACTG 1694
Oy 424 sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr 444
Db 1695 TTTGATTAACAAAAATGAGATGAGAGTACAACTCAGTGGGACAAAGAAACTGAG 1754
Oy 444 gSerLeuArgGluLeuLeuAlaAlaCysTyrrAsnSerGlyLeuArgValAspGlyAlaAl 464
Db 1755 CAGTCTTAAAGATCTTTGAAATGTTTACCAATGGAAGAACTGCTCCCTCAGCAATATTAAT 1814
Oy 464 aLeuTyIleLeuThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleVala 484
Db 1815 TTTCCAGTTTACTAATATGCTGTCGCCCAAGCCAAAGTAATATCAAACTTCTAGTCTT 1874
Oy 484 lArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAlaIle 503
Db 1875 CAGAAGCAATGGTGTCTTGATGATACCACTCACCAACATTACAGAGGCTTACTCATAT 1934
Oy 503 uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluThrPheHisGlnAsnLeu 523
Db 1935 GAACCAAAATGGCTTTCACAAAAATCAGAAATGAGATTTGATTAATGAAGACCTTGG 1994
Oy 523 yHisGlySerPheThrLysIlePheArgGlySerArgGluValValAsp---GlyL 542
Db 1995 CCAGGACACTTTTACAAAGATTTTAAAGCGTACGACAGAGATGAGATACAGCTCA 2054
Oy 542 uThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetG 562
Db 2055 ACTGATGAAGACGAGAGTCTTTAAAGTCTGGATTAAGCACACACAGACATTCACA 2114
Oy 562 uSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyProHisLeuValLeu 582
Db 2115 GTTCTTCTTTGAAGCAGCAAGTATGATGACAGCTTTCACACACATTTGGTTTAA 2174
Oy 582 uHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyLeuG 601
Db 2175 TTATGAGATGTGTCTGTGAGACAGAAATATTTCTGTTACAGAGCTTTGTAATAATTGG 2234
Oy 601 yAlaIleAspMetTyIleuArgLysArgGlyHisLeuValSerAlaSerTyPryLeuG 621
Db 2235 ATACTAGATACATATCTGAAAAAGATTAATAATTTGATTAATATTTATGGAATCTGA 2294
Oy 621 nValThrLysGlnLeuAlaLysAlaLeuAsnTyIleGluAspGlyGlyLeuProHisG 641
Db 2295 AGTTGCTAAACAGTTGGCATGGCCATGCTTTTCTAAGAAAGAAACACCTTATTCATG 2354
Oy 641 yAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGly---AspGlyAsnPr 659
Db 2355 GAATGATATGCGCAAAATATTTCTGCTTATCAGAAAGACAGACAGCAAGAAATCC 2414
Oy 659 oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLe 679

```

---

```

Db 2415 TCCTTCATCAAACTTAGTGCATCTGSCATTAGATTACAGTTTGGCAAGACATTC 2474
Oy 679 uThrAspArgIleProThrPheAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe 699
Db 2475 TCAGGAAGATATACATGAGGATGACCTGAAATGATGAAATTCCTAATAATTAATTT 2534
Oy 699 uGluAlaAspLysTyPheGlyAlaThrThrTyIleThrValPheGlnArgGlyProAl 719
Db 2535 GGCACACACAAATGAGATTTGGTACCCTTGTGGAAATCGACAGTACAGAGATTA 2594
Oy 719 aHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyGlnAspGlnGlyLe 739
Db 2595 ACCCTAAGCTGCTGATCTCAAGAAAGCTCAAAATTTATGATACATGGCATCCGCT 2654
Oy 739 uProAlaLeuLysThrPheGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyAspPr 759
Db 2655 TCCTGCACCAAGTGGGACAGATTAAGCAACCTTATAATTAATGTATGATTAAGAAC 2714
Oy 759 oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779
Db 2715 AGATTTCAGGCTTCTTTCAGAGCCATCATCAGAGATCTTAACAGTTTGTTTACTCCAGA 2774
Oy 779 pTyGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCy 798
Db 2775 TTATGAACTATTAACAGAA---AATGACATGTTCACAAATATGAGATAGGTGCTTGG 2831
Oy 798 sValAlaGluAlaGluLeuTyIleAlaCysGlnAspProAlaIlePheGluGluArgHisLe 818
Db 2832 GTTTCTGCTGTCG---TTTGAAGACCGGAGATCTTACACAGTTTGAAGACACATTT 2885
Oy 818 uLysTyIleSerLeuLeuGlyLysGlyAsnPheGlySerValGlnLeuCysArgTyAs 838
Db 2886 GAATTTTTCAGCAACTTGGCAAGGTAATTTGGAGTGTGGAGATGTCGCCGTATGA 2945
Oy 838 pProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPr 858
Db 2946 CCTCTACAGACACACATGGGAGGTGCTGCTGCTAATAAAGCTTCAGATTAAGTCA 3005
Oy 858 cAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPh 878
Db 3006 AGAGCACTTAAGACCTTGTGAAGGAAATGGAATCTGTAATTCCTAAGCATGACACA 3065
Oy 878 eIleValLysTyIleArgGlyValSerTyGlyProGlyArgGlnSerLeuArgLeuValMe 898
Db 3066 CATTTGAAGTACAGAGGAGTGTGTACAGTGTGTGGGTAAATCTAATAATTAATAT 3125
Oy 898 tGluTyIleuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuH 917
Db 3126 GGAATATTTACCATATGGAAGTTTACGAGACTATCTCAAAAACATTAAGAAAGGATAGA 3185
Oy 917 sThrAspArgLeuLeuPheAlaIleThrGlnIleCysLysGlyMetGluTyIleGlyAl 937
Db 3186 TCACATTAATACTTGCACACTACATCTCAGATATGCAAGGATATGAGATCTTGGTAC 3245
Oy 937 aArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGlnAlaHis 957
Db 3246 AAAAAGTATATCCACAGGATCTGGCAAGCAAGAAATATATTGTGGGAAGACAGAACAG 3305
Oy 957 sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyTyVa 977
Db 3306 AGTTAAATTTGGGATTTTGGCTTAACCAAACTTGGCCACAAGCAAGAAAGATTAATPA 3365
Oy 977 yValArgGluProGlyGlnSerProIlePheThrTyIleAlaProGluSerLeuSerAspAs 997
Db 3366 AGTAAAGAACTGGGAAAGTCCCATATTTGGTATGCTCCAGATCACTACAGAGAG 3425
Oy 997 nIlePheSerArgGlnSerAspValTyrrSerPheGlyValValLeuTyGluLeuPheTh 1017
Db 3426 CAAATTTTCTGGGCTCGATGATGTTGGAGCTTTGGAGAGTGTCTGATGAACCTTTTAC 3485
Oy 1017 rTyCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAC 1037
Db 3486 ATACATTTAGAGAGATTAAGATCCACACGCGGAATTTATGCTATGATTTGGCAATGACA 3545

```

```

QY      1037 gg::gnglyprrobleuyns---argleuleuglulealeualaglulgylarqatgleur 1056
          ::::::::::::::::::::|:::|||||||
Db      3546 ACAAGGACAGATGTCTGTTCCATTGGATGAGAAGCTTTTGAACAATAATGAAAGATTACC 3605
          ::::::::::::::::::::|:::|||||||
QY      1056 opProPrcrhcryspProthrcgluaValngluLeumetGlnLcucySTripAlaPrOgl 1076E
          | ||| ||||| ||:::|:::|:::|:::|:::|
Db      3606 AAGACCAAGATGGAGCCCCACATCATGATCTATATGATCAGTACAGATGCTGTGCAACAATA 3655
          ::::::::::::::::::::|:::|||||||
QY      1076 wProHiasParProAlaPhaelathreUSeSrProGlnInaasPProLeu 1093
          ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      3666 TGTAAATCAACGCCCTTCCTTTAGGGATGTAGCTTCCTTGACAGTGCATCAATA 3717
          ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 8
US-08-097-997A-8
Sequence 8, Application US/08097997A
Patent No. 5728536
GENERAL INFORMATION:
APPLICANT: Ihle, James N.
APPLICANT: Silvenoinen, Olli
APPLICANT: Wittenhub, Bruce A.
APPLICANT: Quelle, Frederick W.
TITLE OF INVENTION: Jak kinases and Regulation of Cytokine Signal
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,997A
FILING DATE: 29-JULY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0370000/SLE/GKT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEO ID NO.: 8
SEQUENCE CHARACTERISTICS:
LENGTH: 3629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 94..3480
US-08-097-997A-8

Alignment Scores:
Pred. No.: 6,47e-234 Length: 3629
Score: 2619.50 Matches: 537
Percent Similarity: 65.50% Conservative: 190
Best Local Similarity: 48.38% Mismatches: 350
Query Match: 44.70% Indels: 33
DB: 1 Gaps: 15

US-09-397-967-16 (1-1099) x US-08-097-997A-8 (1-3629)
QY      4 ProSerCluGluthrProLeuIleProGlnArgSerCysSerLeuSerSerSerCluAla 23
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      145 CCTCATCATGATGAGTAGTGATATTCTCGAAGAGTGCATAATTCGTGTGAACAGATAGAGCCA 204
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

OY	24	GlyAlaLeuHisValLeuLeuProProLysGlyProGlyProProGlnArgLeuSerPhe	43
Db	205	GTCTTCAAGTAGATGTCACATTCTTTGGGCAAGTGAAGAAGATGTAAAGTTT	264
OY	44	SerPheGlnAspTryLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIle	63
Db	265	CCAAGTGAGAGATGTGTGCAGAAAGAAATTTGTGTGGCGCTTCTTAACCTGTGTATTT	324
OY	64	LeuProValTrpHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro	83
Db	325	ACGGCTGTGATCATATATATGTTGGCTTAATAGTGAACCGAAAGATCTGTACACC	384
OY	84	ProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTrpArgLeuArg	103
Db	385	CCCAATCATGTCTTCCACATAGACAGATCACACGAGCATGACATACTCTACAGGTAAAG	444
OY	104	PheTyrrPheProAspTrpPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg	121
Db	445	TTCCTACTCCCTCATTTGGTACTAGTGCAGCAGACAAGCTACAGATACGGATGTGCC	504
OY	122	LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGlnHisIleuPheAlaHis	141
Db	505	CGTGGGCGTCAAGCTCTCTGCTTGATGACCTTTGTATGTCTTACCTTTTGTCTCAAGCG	564
OY	142	ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetCysGluGlnGly	161
Db	565	CGCGATCATTTTTGTCACGGATGGATTAAGATACCTGTGACATGAAACCTCAGGAAG	624
OY	162	PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro	181
Db	625	TGCTGTGGAGGGGGGTTTGAAGCATGATGATGAGATGTGCTAACGAAAGAACCGACTCCA	684
OY	182	GlyIleuLeuLeuLysThrValSerTrpLysValCysLeuProProSerLeuArgAspAl	201
Db	685	CTGGCTGTCTATAACTCTGTACACTCAAGACATTTCTTACCAAAGTGGCTTGACCGAG	744
OY	202	IleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAlaLeu	221
Db	745	ATCCAAACATATCATCATTTTAACCCGGAAGCGAATCAGATACGATTTCCGAGATTCTT	804
OY	222	LeuProCysGlyArgLeuProGlyArgProTrpArgAlaLeuMetAlaLysTrpIleLeuasp	241
Db	805	CAGCAATTCACATCATTTAAAGCCACTGCCAGAGACCTTAAACCTTAAGTATCTTAAAC	864
OY	242	LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal-----	256
Db	865	CTGGAACCCCTCAGCTGCTGCTCTACACAGAAACGTTTGAATGAATAAGATCTCAGAG	924
OY	257	GlyLeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle	276
Db	925	GGT---CCTTCAGGTGAGAGATTTTGCACACCATTATATAGTCGAAGAACCGTGAATTT	981
OY	277	ProTrpSerSer-----AsnAspGluLeuPheGlnThr	287
Db	982	CAGTGTTCAGAGAGAAACATTAAGAAAGAGACCTGCAGAACAGACAGCTACAGTTA	1041
OY	288	PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValCysPro	307
Db	1042	TATTTGATTTTCCCTGATATATATTTGATGTGCAGATTTAAACCAAGAACCCAG---GAATGC	1098
OY	308	AlaGlyLeuHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGlnValGlu	327
Db	1099	TCAATTAAGAACTGAATTTGAATCTGTCCATTAACAAAGATGCTAAAGTTTGGACGTAGA	1158
OY	328	PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspCylTyrrPheArgLeu	347
Db	1159	CTTAGCTCATTTAAAGAAAGCTGTGATTCGCTCATTTGATTTAGACGGGATATTCAGACTA	1218
OY	348	IleCysAspSerArgHisTrpTrpPheCysLysGlyValAlaProProArgLeuLeuGlnIu	367
Db	1219	ACTGCGATGGCCACCACTTACTCTGTCAAAAGAGGTGGCTCCCCACGCTGTCTCAGAAC	1278
OY	368	GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAla	387

```
Db 1279 ATACAGACAACCTCCACGGCCCAATATCATGATTTTGGCATTTAGCAAACTAAGAG 1338
OY 388 AlAglyserLeuProGlyThrTlleuAArgSerProGlnAspTyrAspSerPhe 407
Db 1339 GCGGATTAACGAGCTGATATGCTAGACGACCCCTAAGACCTTCAACAAATAC 1398
OY 408 LeuLeuThralAcysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArg 427
Db 1399 TTTTGTACCTTGTCTGTGACCGGAAAAATGTCATGATATTAACACTGTTTATTCAG 1458
OY 428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447
Db 1459 AAGATGAGATGAGATGAGATACACCTCAGCGGACTAAGAGAACTTCAGTAACCTTAAG 1518
OY 448 GluLeuLeuAlaAlaCysTTrpAsnSerGlyLeuArgValAspGlyAlaIleuTyrLeu 467
Db 1519 GACCTTTTGAATTCCTACACAGATGAAACGTGCGCTCAGACATATCATCTTCACGTTT 1578
OY 468 ThrSerCysAlaProArgProLysGlyLysSerAsnLeuIleValAlaArgGly 487
Db 1579 ACCAAATGCTGCCCCCAACCAACCAATATCAACCTTCGCTTCAGA----- 1632
OY 488 CysAsnProAlaProAlaProGlyCysSerProSerCysAla-----LeuThr 504
Db 1633 ACAATGCTATTTCTGATGTTCAATCTCACCAACATTAACAGGACATATATATGTAAT 1692
OY 505 GlnLeuSerPheHisThrIleProThrsPheSerLeuGlnTyrPheHisGlnAsnLeuGlyHis 524
Db 1693 CAAATGCTTTCACAAATAGAGATGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1752
OY 525 GlySerPheThrLysIlePheArgGlySerArgArgGlyValAlaAsp--GlyGlnThr 543
Db 1753 GGTACTTTTACAAAATTTTAAAGTGTAGAGAGAGAGAGAGAGAGATTAATGCAACTG 1812
OY 544 HisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSer 563
Db 1813 CACAAACGGAGATGCTTTTCAAAAGTCTAGATTAAGCACATAGGAACATTAATTCAGAGCT 1872
OY 564 PheLeuGlnAlaIleSerLeuMetSerGlnValSerTyrProHisIleValLeuLeuHis 583
Db 1873 TTTCTCGAAGCAGCAGCATATGATGATGCTTTCTCAAGCATTTGGTTTGAATATAT 1932
OY 584 GlyValCysMetAlaGly--AspSerIleMetValGlnIlePheValTyrLeuGlyAla 602
Db 1933 GGTGCTGTGCTGTGAGAGAGAGAACTCTGTTCAAGATTTGTAAATTTGGATCA 1992
OY 603 IleAspMetTyrLeuArgGlyArgGlyHisIleuValSerAlaSerTrpLysLeuGlnVal 622
Db 1993 CTGATACATACCTGAGAGAGAAACAAATAATTCATTAATTAATTAATTAATTAATTAAT 2052
OY 623 ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnAspLysGlyLeuProHisGlyAsn 642
Db 2053 GCTAAGCAGTGGGATGGGCGCATGATTTTCTAGAGAAAAATCCCTTATTCATGAGGAT 2112
OY 643 ValSerAlaArgLysValLeuLeuAlaArgGlnGlyGly-----AspGlyAsnProPro 660
Db 2113 GTGTGCTGCTAAATAATCTCTTATCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2172
OY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGlnMetLeuThr 680
Db 2173 TTTCAATCAAACTTAGATCCCTGAGATTAGCATTTACAGTTTACGAAAGAGACATTTCTTAG 2232
OY 681 AspArgIleProTrpValAlaProGlnCysLeuGlnGlnIleValIleThrLeuCysLeu 700
Db 2233 GAGGAAATACCATGGGTACCTCCCTGAATGCATTTGAGAAATCTTAAATAATCTAATCTGACA 2292
OY 701 AlaAspLysTrpGlyPheGlyAlaThrThrTrpIleValPheGlnArgGlyProAlaHis 720
Db 2293 ACAGACAGATGAGCTTGAGGACCACTCTGGAGATCTCAGTGGAGAGATTAAGCCC 2352
OY 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGlnAspGlnGlnLeuPro 740
```

---

```
Db 2353 CTGAGTCTCTGATTTCTCAAGAAACCTGAGTCTTATGAGATTAAGATCACTTCTCT 2412
OY 741 AlaLeuLysTrpPheArgLysLeuAlaGlyLeuIlePheGlnCysMetAlaTyrAspProGly 760
Db 2413 GCACCAAGTGGAGACAGATTGACAAACCTTATTAATTAATTAATTAATTAATTAATTAATTA 2472
OY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780
Db 2473 TTTAGGCTGCTTTCAGAGCTGTCATCCGTATCTTAAACGCTGTTTACCCAGATTAAT 2532
OY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArg--AspGluLeuCysVal 799
Db 2533 GAACCTTAAACGAA---AATGACATCTACCAACATGAAATAGGTGCCCTTAGGGTTT 2589
OY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGlnIleArgHisLeuLys 819
Db 2590 TCTGGTCT-----TTTGAAGACAGGAGACCCCTACACAGTTTGAAGACACACTTGAAG 2643
OY 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPro 839
Db 2644 TTTCTACAGCAGCTTGGCAAAAGTAACTCGGAGTGTGGAGATGTCCCTATGACCCG 2703
OY 840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859
Db 2704 CTGAGAGCAACACTGGCGAGGTGCTGCTGTGAAGAACTCCGACACAGCATGAAAG 2763
OY 860 GlnGlnArgAspPheGlnArgGlyIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879
Db 2764 CACCTCCGACACTTTGAGAGGAGATGACAGATCCCTGAATCTTGACAGATCAACATC 2823
OY 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899
Db 2824 GTCAAGTACAAAGAGAGTGTCTACAGTGGCGGTGGCGCAACCTTAATTAATTAATTAATTA 2883
OY 900 TyrLeuProSerGlyCysLeuArgArgAspLeuLeuIleArgHisArg--GlyLeuHisThr 918
Db 2884 TATTTACCATATGGAAGTTTACGAGATATCTCCAAAAACCTTAAGAAAGAGATGATGATC 2943
OY 919 AspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlnMetGluTyrLeuGlyAlaArg 938
Db 2944 AAAAAGCTTCTTCATACATACATCTCAGATATGCAAGGAGCATGATATCTTGTAACAAA 3003
OY 939 ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVal 958
Db 3004 AGGTATATCCACAGAGGAGCTTGGCAACAAAGCATTTGCGGAAATGAGAACAGAGGCTT 3063
OY 959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValAl 978
Db 3064 AAAATAGAGAGCTTGCGATTACCAAAAGTCTTGCCGACAGCAAAAGATTAATCAAAAGTA 3123
OY 979 ArgGluProGlyGlnSerProIlePheTrpTyrAlaProGlnSerLeuSerAspAsnIle 998
Db 3124 AAGAGCGCAGGGAAGCCCATATTTCTGTAGCCACCTGAATCTTGACGAGAGAGAGAG 3183
OY 999 PheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGlnLeuPheThrTyr 1018
Db 3184 TTTCTGTGCTCAGATGAGTGTGAGCTTTGAGTGTCTTATCGAAGCTTTTACATAC 3243
OY 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGlnArgGlu 1038
Db 3244 ATCGAAGAGATTAAGATCCACCCCTGGAATTTATGCAATGATTTGGCAATGATTAACAA 3303
OY 1039 GlyProProLeuCys--ArgLeuLeuGlnLeuAlaGlnIleArgValArgValLeuProPro 1057
Db 3304 GGGCAAAATGATTTGTTCCATTTTATGATAGACTACTGAAGAGCAACGAGATTTGCCAAG 3363
OY 1058 ProProThrCysProThrGlnValGlnIleuMetGlnLeuCysTrpAlaProGlnPro 1077
Db 3364 CCAGAGAGATCCCAAGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3423
OY 1078 HisAspArgProAlaPheAlaThrLeuSer 1087
Db 3424 AGCCAGGCTCCCTCTTACGAGGACCTTTCG 3453
```



1339 GCGGATACAGACTGACACTATATGCTACGATGACGCCCTTAAGCCTTCAACAATAC 1398  
QY LeuleuthralacysValGlnThrProLeuGlyProaspTyrLysGlyCysLeuIlearg 427  
1399 TTCTGACCTTTGCTGTTGAGCGGAGAAATGTCATTAATATAAACCTGTTGATTACG 1458  
QY 428 GlnAspProserGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuarg 447  
1459 AAGATAGAGATGAGATGACATCAACCTCAGCGGACTAAGAGGAACTTCAGTAACCTTAAG 1518  
QY 448 GlnLeuLeuAlaIacysTyrPasnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 467  
1519 GACCTTTTGATTTGCTACGACGATGGAACCTGCGCTCAGACGATACATCTTCCACTTT 1578  
QY 468 ThrSerCysAlaProArgProLysGlyLysSerGlnLeuIleValIleArgArgGly 487  
1579 ACCAAATGCTGCCCCCAAGGCCAAAGATTAATCAAACTTCGCTCCTCAGA----- 1632  
QY 488 CysAsnProAlaProAlaProGlyCysSerProSerCysAla-----LeuThr 504  
1633 ACAATGCTATTTCTGATGTTGATGCTCAGACATTCACAGGCTAATATGTAAT 1692  
QY 505 GlnLeuSerPheHisThrIleProThrAspSerLeuGlnTyrPheHisGlnLeuGlyHis 524  
1693 CAATGCTGTTTCCAAATCAGACGATGAGATTAATATTAATGAAGCTTGCCCAA 1752  
QY 525 GlySerPheThrLysIlePheArgGlySerArgArgGlyValAlaIleArgGlyThr 543  
1753 GGTCTTTTACAAAAATTTTAAAGGTGTAAAGAGAGAGATGAGATTTATGCTCAACTG 1812  
QY 544 HisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnGlySer 563  
1813 CACAAACGAGAGCTTTTAAAGCTTCAATGAACACACATAGACACTATTCAGAGCT 1872  
QY 564 PheLeuGlnAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuHis 583  
1873 TTCTTCAGAGCAGACACATGATGAGCTTTCACAAAGCATTTGGTTTAAATAT 1932  
QY 584 GlyValLysMetIleArgLysAspSerIleMetValGlnGlnPheValTyrLeuGlyVal 602  
1933 GGTCTGCTGCTGCTGGAGAGAGACATTTGCTTCAAGATTTGTTAAATTTGGATCA 1992  
QY 603 IleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrLysLeuGlnVal 622  
1993 CTGATACATCTGAGAGAACAAATTCATTAATATATATGGAATCTTGAGAT 2052  
QY 623 ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlnLysPheGlyLeuProHisGlyAsn 642  
2053 GCTAAGCAGTTGGCATGGCCATGCTATTTCTAAGAGAAATCCCTTATTCATGGGAT 2112  
QY 643 ValSerAlaArgLysValLeuLeuAlaArgLysGlyLys-----AspGlyAsnProPro 660  
2113 GTGCTGCTAAATATTCCTCTTATCAGAGACAGACAGACAGACAGGAGGAGACCCACT 2172  
QY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGlnMetLeuThr 680  
2173 TTCTCAAACTTAGTATCCGCGATTTAGCATTTACAGTTTACGAGAGGACATTTCTCAG 2232  
QY 681 AspArgLysLeuProTyrAlaAlaProGlnLysLeuGlnGlnIleThrLeuCysLeuGlu 700  
2233 GAGGAAATACCATGGTACCTCTGAATGCAATGAGAAATCCTAAATCTCAATCTGSCA 2292  
QY 701 AlaAspLysTyrPheGlyAlaIleThrThrArgValPheGlnArgGlyProAlaHis 720  
2293 ACAGACAGATGGAGCTTGGGACCACTCTGGGAGATCTGACAGGAGAGAGATAGGCC 2352  
QY 721 IleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGlnAspGlnGlyLeuPro 740  
2353 CTGAGTGTCTGAGATTCCAAAGAAAGCTGCAAGTTCTATGAAGATTAACATCAGCTTCT 2412  
QY 741 AlaLeuLysTyrPheGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760

2413 GCACCCAGTGGACAGAGTAGCAAACTTATAAATATTCATGAGCATATGACCCAGAT 2472  
QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780  
2473 TTCAGGCTCTCTTCAAGAGTGCATCCGATCTTAACAGCCTGTTTACTCCAGATTAT 2532  
QY 781 GlnLeuLeuSerAspProThrProGlyIleProSerProArgLys-----AspGlyLeuCysVal 799  
2533 GAACCTACTACAGAA-----AATGACATGCTACCAAACTAGAAATAGTGCCCTAGGGTTT 2589  
QY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGlnIleThrLys 819  
2590 TCTGCTGCT-----TTGAAGACAGGAGACCTTACACAGTGTGAAGAGACACTTGAG 2643  
QY 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGlnLeuCysArgTyrAspPro 839  
2644 TTCTTACAGCAGCTTGGCAAGGTAACCTTGGAGTGTGAGATGTGCCCTTATGACCG 2703  
QY 840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859  
2704 CTGAGAGCAACACTGGCGAGGTGCTGCTGTAAGAACTCCAGCAGACACTGAAGAG 2763  
QY 860 GlnGlnArgAspPheGlnArgGlnIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879  
2764 CACCTCGAGACTTGGAGAGGAGATGAGATCCCTGAATCCTTGACAGCATGCAACATC 2823  
QY 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899  
2824 GTCAGATCAAGAGAGCTGCTACAGTGGCGGTGCGGCAACCTAAGATTAATATGAA 2883  
QY 900 TyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg-----GlyLeuHisThr 918  
2884 TATTTACCAATGAGAGCTTACGAGACTATCTCCAAATGAAGAAGCAGATGATCAC 2943  
QY 919 AspArgLeuLeuLeuPheAlaTyrGlnIleCysLysGlnMetGlnTyrLeuGlyAlaArg 938  
2944 AAAAATCTCTCATACATCAGATCTCAGATTCAGAGGCTGGAATTTCTTGACAAAA 3003  
QY 939 ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGlnAlaHisVal 958  
3004 AGTATATTCACAGAGGAGCTGGCAACAGAACATATTTGGTGGAAATGAGAACAGCGTT 3063  
QY 959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal 978  
3064 AAAATAGAGAGCTTGGATTAACCAAGCTTGGCGAGGACAAAGATCTTACAAAGTA 3123  
QY 979 ArgGlnProGlnGlnSerProIlePheTyrTyrAlaProGlnSerLeuSerAspAsnIle 998  
3124 AAGAGCGCAGGAGAAACCCCATTTCTGTAGCCACCTGAATCCTTGACGAGAGCAG 3183  
QY 999 PheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGlnLeuPheThrTyr 1018  
3184 TTCTTGGCGCTCAGATGAGTGTGAGCTTGGAGCTTGTATACGAACCTTTCACATAC 3243  
QY 1019 CysAspLysSerCysSerProSerAlaGlnPheLeuArgMetMetGlyProGlnLysGlu 1038  
3244 ATCAGAGAGATTAAGTCCACCCGCTGGAATTTATGCGAATGATTTGCGATGATTAACAA 3303  
QY 1039 GlyProProLeuCys-----ArgLeuLeuGlnLeuLeuAlaGlnLysArgArgLeuPro 1057  
3304 GGGCAATGATTTGTTCCATTTGATAGAGCTACTGAAGACAGCAAGGAAATTCACAAAG 3363  
QY 1058 ProProThrCysProThrGlnValGlnGlnLeuMetGlnLeuCysTyrPheAlaProGlnPro 1077  
3364 CCAGAGAGATCCCAAGATTAAGATTTATGTATCATGTACAGAGTGTGGAACACAAATGTG 3423  
QY 1078 HisAspArgProAlaPheAlaThrLeuSer 1087  
3424 AGCCAGGCTCCCTTCAGGACCTTTCG 3453  
Db 3424 AGCCAGGCTCCCTTCAGGACCTTTCG 3453  
RESULT 10  
US-08-946-994-8  
; Sequence 8, Application US/08946994



Patent No. 6210654  
GENERAL INFORMATION:  
APPLICANT: Ihle, James N.  
APPLICANT: Silvenoinen, Olli  
APPLICANT: Wittenh, Bruce A.  
APPLICANT: Queller, Frederick W.  
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/946,994  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,574  
FILING DATE: 18-JUN-1996  
APPLICATION NUMBER: 08/282,012  
FILING DATE: 29-JUL-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/097,997  
FILING DATE: 23-JUL-1993  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/118,968  
FILING DATE: 09-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0370002/SLE/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3629 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 94..3480  
US-08-946-994-8  
Alignment Scores:  
Pred. No.: 6.47e-234 Length: 3629  
Score: 2619.50 Matches: 537  
Percent Similarity: 65.50% Conservative: 190  
Best Local Similarity: 48.38% Mismatches: 350  
Query Match: 44.70% Indels: 33  
DB: 4 Gaps: 15  
US-09-397-967-16 (1-1099) x US-08-946-994-8 (1-3629)  
QY 4 ProsergluigiutnrproleuileproglinarseryserserleuuserSerSerGluAla 23  
Db 145 CCTGTACTCATCGATGCGATATTCCTCGGAAGTGTCTAATTCGTGAAGCAGATAGACCA 204  
QY 24 GAlvalaleuHlsvallleuLeuProProArgGlyProGlyProProGlnAtrgleuSerPhe 43  
Db 205 GTCCCTTCAAGTGTATCTGTACATTTCTTGGCAGACCTGAGAGAGATGTCTGAGCTTT 264  
QY 44 SerPheGlyAspTyrLeuAlaGluAspLeuGlyValAlaGlnAlaAlaGlyAlaGlyAla 63

Db 265 CCAAGTCAGAGATGATGTCAGAGAAATTTGTGTGGCTGCTTCAAGCTTGTGGTATT 324  
QY 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTyrPhePro 83  
Db 325 ACCGCTGTATCATTAATAATGTTTGTGCTTAAGTGAACGGAAGATGTGGTACCA 384  
QY 84 ProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg 103  
Db 385 CCCAATCATGTCTTCCACATAGACAGTCACACAGCAGCATGACATCTTACAGGATTAAG 444  
QY 104 PheTyrPheProAspTyrPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg 121  
Db 445 TTCTACTTCCCTCATTTGTACTAGTACTGTGTCAGCAGCAGAACCTCAGATCAGAGTGC 504  
QY 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141  
Db 505 CGTGGGCTGAAGCTCCCTGCTGTGATGACTTTGTCTATGCTTCTTGTGGCTCAGTGC 564  
QY 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyu 161  
Db 565 CGGCATGATTTTGTTCACGAGATGATGAAGTACCTGACTCATGAACTCAGAGAGAG 624  
QY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnPro 181  
Db 625 TGTCTGGATGCGCGGTGTGTAGACATGTGAGATTAAGTAAAGAGAAAGACACTCCA 684  
QY 182 GlyLeuLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201  
Db 685 CTGCTGTCTTAACCTGTGACCTACAGACATCTTCCAAAGTCCGTTCCAGCAG 744  
QY 202 IleGlnGlyGlnAspPheValThrArgArgArgIleArgArgThrValValLeuAlaLeu 221  
Db 745 ATCCAGACTATACATTTTAACCCGAGACGATCCGATTCAGATTTCCAGATTCATT 804  
QY 222 LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp 241  
Db 805 CAGCAATTCAGTCAATTAAGCCACGACGACGACCTAAACTTAAGTATCTTAAAC 864  
QY 242 LeuGluArgLeuHisProAlaIleThrThGluThrPheArgVal----- 256  
Db 865 CTGGAACCTGACGCTGCTTCTACACAGAACAGTGTAAAGATTCGACAG 924  
QY 257 GlyLeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle 276  
Db 925 GGT---CCTTCAGGTAGAGATTTTGCACCATTTATTAATACGTAACGGTGAATT 981  
QY 277 ProTyrPheSer-----AsnAspGluLeuPheGlnThr 287  
Db 982 CAGTGTCAAGAGGAAACATAAGAAAGTGAACACAGACAGACAGACAGACAGCTTA 1041  
QY 288 PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro 307  
Db 1042 TATTGTGATTTCCCTGATATTATGATGTCAGATTAAGCAAGAACAGC---GATATC 1098  
QY 308 AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGlnAlaGlu 327  
Db 1099 TCAAATGAAGTGAATGTAACTGTCAATAAACAAGATGTAAGTTTGGAGATAGAA 1158  
QY 328 PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347  
Db 1159 CTTAGCTCATTAATAAACAACCTTGTCTATTCGTGTATTAATGAGCGGATTAACAGACTA 1218  
QY 348 IleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGlu 367  
Db 1219 ACTGGGATGCCACACATTAACCTGTGCAAGAGGTGGCTCCCAAGTGTCCAGAAC 1278  
QY 368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAla 387  
Db 1279 ATACACAGCAACTGCCAGCCCAATATCAATGATTTTGGCATTAAGCAACTTAAGAG 1338  
QY 388 AlaGlySerLeuProGlyThrTyrIleLeuArgAspSerProGlnAspTyrAspPhe 407

Db 1339 GCGGGTAACAGACTGGACTATATGTACGATCGACCCCTAAGAGCTTCAACAATATAC 1398  
OY LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrIleuSyluysLeuArg 427  
Db 1399 TTTCTGACCTTCTCTGTCGCGAGAAATGTCATTGAATTAACAGCTTTGATTACG 1458  
OY 428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447  
Db 1459 AAGATAGATAGATGAGAAATACACCTCAGCGGACATAAGAGAACTTCAGTAACCTTAAG 1518  
OY 448 GlnLeuLeuAlaIleCysTyrPasnSerGlyLeuArgValAspGlyAlaIleuThrLeu 467  
Db 1519 GACCTTTGAATTCCTCAGCATGGAACCTGCGCTCAGACAGATATCATCTTCCAGTTT 1578  
OY 468 ThrSerCysAlaProArgProGlySerGlnuysSerAsnLeuIleValAlaArgArgGly 487  
Db 1579 ACCAAATGCTGCCCCCAAGGCCAAAGATTAATCAACCTTCTGCTCACA----- 1632  
OY 488 CysAsnProAlaProAlaProGlyCysSerProSerCysCysala-----LeuThr 504  
Db 1633 ACAATGATATTTCTGATGTTCAAGATCTCACACATTAACAGAGCATTAATATGTAAT 1692  
OY 505 GlnLeuSerPheHisThrIleProThrAspSerLeuGlnThrPheHisGlnuysLeuHis 524  
Db 1693 CAAATGCTTTCACAAATCAGCAATGCAAGATTATTAATTAAGAACTTGGCCAA 1752  
OY 525 GlySerPheThrIlePheArgGlySerArgArgGluValAlaAsp---GlyGluThr 543  
Db 1753 GGTACTTTTACAAAAATTTTAAAGGTGTAAAGAGAGAGATGAGATTTATGTCACACTG 1812  
OY 544 HisAspSerGluValLeuLeuIleuysValMetAspSerArgHisArgAsnCysMetGluSer 563  
Db 1813 CACAAAGCGGAGTCTTTTGAAGTCTTACATTAAGCACATAGCACTATTCAGACTCT 1872  
OY 564 PheLeuGlnAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHis 583  
Db 1873 TTCCTGCAAGCAGCAAGCATGATGATGCTTCTCACAAGCATTTGGTTTGAATAT 1932  
OY 584 GlnValCysMetAlaGly---AspSerIleMetValGlnGluPheValTyrLeuGlyAla 602  
Db 1933 GGTGCTGTGCTGTGAGAGAGAGAACATTCGGTTCAGAAATTTGTAAATTTGGATCA 1992  
OY 603 IleAspMetTyrLeuArgIleValLeuValHisLeuValSerAlaSerTyrPheLeuGlnVal 622  
Db 1993 CTGGATACATACCTGACAGAGAACAAATAATCCATAATATTATGAAACTTGGAGATG 2052  
OY 623 ThrIysGlnLeuAlaIleTyrAlaLeuAsnTyrLeuGluAspIleuysGlyLeuProHisGlyAsn 642  
Db 2053 GCTAAGCAGTTGGCATGGCCCATGTCATTTCTTGAAGAAATCCCTTATTCTATGGGAT 2112  
OY 643 ValSerAlaArgIleValLeuLeuAlaArgGluGly-----AspGlyAsnProPro 660  
Db 2113 GGTGTGCTTAAATAATCTCTGCTTATCAGAGAAAGACAGAGAAAGGAGAACCCACT 2172  
OY 661 PheIleIleuysLeuSerAspProGlyValSerProThrValLeuSerLeuGlnMetLeuThr 680  
Db 2173 TTCATCAAACTAGTATGATCTGTCATTAACATTACAGTTTCGAGAGACATTTCTTACG 2232  
OY 681 AspArgIleProThrValAlaProGluCysLeuGlnGluAlaGlnThrCysLeuGln 700  
Db 2233 GAGAGATACCATGAGTCTGCTCTGATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2292  
OY 701 AlaAspIleTyrPheGlyAlaIleThrThrTyrPheGluValPheGlnArgIleProAlaHis 720  
Db 2293 ACAGACAG 2352  
OY 721 IleThrSerLeuGlnProAlaIleuysLeuysPheTyrGluAspGlnGlnLeuPro 740  
Db 2353 CTGAGTGTCTCTGATTCCTCAAGAGAAAGCTGCACTTCTATGAGATTAAGATCAGCTTCT 2412  
OY 741 AlaLeuIleThrThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760  
Db 2413 GCACCCAGAGTGACAGAGATTAGCAAACTTATAATATTTGATGAGATATGAGCCAGAT 2472

OY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780  
Db 2473 TTTACGCGCTGTCTTCAGAGCTGATCCGTGATCTTAACAGCCCTGTTTACTCCAGATAT 2532  
OY 781 GlnLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCysVal 799  
Db 2533 GAACTACTAACAGAA---AATGACATGCTTACCAAAACATGACAAATAGAGTCCCTAGGCTT 2589  
OY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluAlaGlnHisLeuys 819  
Db 2590 TCTGCTGCT-----TTTGAAGACACAGGAGCCCTACACAGATTGGAAGAGACACTTGAAG 2643  
OY 820 TyrIleSerLeuLeuGlnIleuysGlnAsnPheGlySerValGlnLeuCysArgTyrAspPro 839  
Db 2644 TTTTACACAGCTTGGCAAGAGTAACTTCCGAGATGAGAGATGCGGCTATGACCCG 2703  
OY 840 LeuGlyAspAsnThrGlyProLeuValAlaValIleuysGlnLeuGlnHisSerValProAsp 859  
Db 2704 CTGACAGACACACACTGCGAGGTGCTGCTGTAAGAAATCCACACAGCAGCATGAGAG 2763  
OY 860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuysAlaLeuHisSerAspPheIle 879  
Db 2764 CACCTCCGAGCTTGTAGAGGAGATCCAGATCTCGAATCTTCCAGCATGACAAATC 2823  
OY 880 ValIleTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899  
Db 2824 GTCAACTACAGAGAGATGCTGCTACAGTCCGCTGCGCGCAACCTTAATTAATTTGGA 2883  
OY 900 TyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuIleThr 918  
Db 2884 TATTTACATGAGAGAGTGTACAGACTTCTCCAAACATTAAGAAAGCATGATATC 2943  
OY 919 AspArgLeuLeuLeuPheAlaThrGlnIleCysIleuysGlyMetGluTyrLeuGlyAlaArg 938  
Db 2944 AAAAATCTTCAATACATACATCTCAGATATGCAAGGAGCATGGAATCTTGGTCAAA 3003  
OY 939 ArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValGlnSerGluAlaHisVal 958  
Db 3004 AGGTATATCCACAGGAGCTGCGCACACAGAACATTTGTTGGAATATGAAACAGGTT 3063  
OY 959 LysIleAlaAspPheGlyLeuAlaIleuysLeuLeuProLeuGlyLysAspTyrTyrValVal 978  
Db 3064 AAAATAGAGACCTTGGATTAACCAAAAGCTTCCACAGACAAAGATTAACAAGTA 3123  
OY 979 ArgGluProGlyGlnSerProIlePheTyrTyrAlaProGlySerLeuSerAsnIle 998  
Db 3124 AAGGAGCCAGGAGGAAAGCCCATATTTCTGTACGCACTGTAATCTTGAAGGAGCAAG 3183  
OY 999 PheSerArgGlnSerAspValIlePhePheGlyValValLeuTyrGlnLeuPheThrTyr 1018  
Db 3184 TTTCTGTGCTTCAAGATGTGTGAGCTTTGAGTGGTTCTATACGAACTTTTACATAC 3243  
OY 1019 CysAspIleSerCysSerProSerAlaGluPheLeuArgMetIleGlyProGlnAlaArgGlu 1038  
Db 3244 ATCAGAGAAAGTAAAGCTCCACCCGTTGAATTTATGGAATGATGGAATGATGGAATG 3303  
OY 1039 GlyProProLeuys---ArgLeuLeuGlnLeuLeuAlaGlnIleArgLeuProPro 1057  
Db 3304 GGGCAATGATGTGCTTCATTTATAGACTACTGAAAGCAAGCAAGATGCGCAAG 3363  
OY 1058 ProProThrCysProThrGluValGlnGluLeuMetGlnLeuysTyrPheProGluPro 1077  
Db 3364 CCAGAGAGATCCCAAGATGATTAATGATGATGATGATGATGATGATGATGATGATGATG 3423  
OY 1078 HisAspArgProAlaPheAlaIleThrLeuSer 1087  
Db 3424 AGCCAGCTGCTCTCTTACGAGCCTTTCG 3453

RESULT 11  
US-08-446-038B-2.  
; Sequence 2, Application US/08446038B  
; Patent No. 5658791

```

: GENERAL INFORMATION:
: APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;
: APPLICANT: Harpur, Ailsa
: TITLE OF INVENTION: No. 5658791el Protein Tyrosine Kinase
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Felte & Lynch
: STREET: 805 Third Avenue
: CITY: New York City
: STATE: New York
: COUNTRY: USA
: ZIP: 10022
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 Inch, 360 kb storage
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: Wordperfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/446.038B
: FILING DATE: 19-MAY-1995
: CLASSIFICATION: 424
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/064,067
: FILING DATE: 30-Jun-1993
: APPLICATION NUMBER: PCT/US91/08889
: FILING DATE: 26-No. 5658791-1991
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: Australian PK3594/90
: FILING DATE: 28-No. 5658791-1990
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: Australian 88229/91
: FILING DATE: 27-No. 5658791-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Hanson, No. 5658791man D.
: REGISTRATION NUMBER: 30,946
: REFERENCE/DOCKET NUMBER: LUD 5244
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-688-9200
: TELEFAX: 212-638-3884
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3495 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: nucleic acid
: US-08-446-038B-2

```

Alignment Scores:

Pred. No.:	9,436-210	Length:	3495
Score:	2360.00	Matches:	489
Percent Similarity:	65.89%	Conservative:	164
Best Local Similarity:	49.34%	Mismatches:	300
Query Match:	1	Indels:	38
DB:	1	Gaps:	15

```

US-09-397-967-16 (1-1099) x US-08-446-038B-2 (1-3495)
QY 128 IIELEASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 147
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1 CTGCTTGATGACCTTTCTCATCTTACCTTCCCTCAGTGGCGGATGATTTTGTTCAC 60
QY 148 GLEUASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 167
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 61 GGATGATGAAGAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 168 LEUASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 187
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 121 TTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 188 VALSERGLYSERASPLEUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEU 207
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 181 GTCAGCTACAGACATTTCTTACCAAGTGCCTTCGAGCGAAGATCCAGACATTCACATT 240

```

```

QY 208 VALTHIRARGARGLEUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEU 227
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 241 TTAAACCGGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 228 PROGLIYASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 247
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 301 AAGCGACCTGCGACGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 248 ALAALATHTHGLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 262
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 361 GCTTCTACACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 417
QY 263 GLUGLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 280
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 418 GAGATTTTGGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 477
QY 281 -----ASNASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 293
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 478 CATAGGAAAGTACGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 537
QY 294 IIEVALASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 313
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 538 ATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 594
QY 314 VALTHIRVALTHIRARGARGLEUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEU 333
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 595 GTGACCGCTGCACAGACGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 654
QY 334 ALAALASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 353
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 655 GCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 714
QY 354 TYRPHESGLYSGLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 373
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 715 TACCTTGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
QY 374 GLYPROIETHIRLEUASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEU 393
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 775 GCGCCAAATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 834
QY 394 THIRYTHIRLEUASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 413
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 835 CTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 894
QY 414 GINTHIRPROLEUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 433
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 895 GACCGGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954
QY 434 PHEASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 453
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 955 TACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1014
QY 454 TTPASNSERGLYSGLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 473
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1015 CAGATGGAAGTGTGCGCTGACGACGATGATGATGATGATGATGATGATGATGATGAT 1074
QY 474 ARGPROLYSGLYSGLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 493
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1075 AAGCGGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1128
QY 494 PROGLIYASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 510
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1129 GTTCACCTGCTACACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188
QY 511 IIEPROTHIRASPLUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEUVALSER 530
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1189 ATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1248
QY 531 PHEARGGLYSERARGARGLEUHNISVALLEUHNISLEUPHEALAGLNHSISRGSERASPLEU 549
: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 1249 TTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308

```

QY 550 LeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaLysSer 569  
 |||||  
 Db 1309 TTGAAGTCTAGATAAGACATAGAAACATATTCAGAGCTTTCTTCTTGAAGCAGCAAGC 1368  
 QY 570 LeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGly 589  
 |||||  
 Db 1369 ATGATGAGTACAGCTTCTCAGCAACGATTTGGTTTGAATTTGAGATGATGCTGTGGA 1428  
 QY 590 ---AspSerIleMetValGlnGluPheValTyrLeuGlnAlaIleAspMetTyrLeuArg 608  
 |||||  
 Db 1429 GAGAGAGAACATTTTGGTTCAGAGACTTTGTAATAATTGGATCAGTACCTACTACTGAG 1488  
 QY 609 LysArgGlnHisLeuValSerAlaSerTyrPheLeuGlnValThrLysGlnLeuAlaTyr 628  
 |||||  
 Db 1489 AAGAACAAATTTCTATTAATATATTTATGAACTGTGAGAGCGGAGCAAGCACTTGCGATG 1548  
 QY 629 AlaLeuAsnTyrLeuGlnValAspLysGlyLeuProHisGlnValSerAlaArgVal 648  
 |||||  
 Db 1549 GCCATGCACTTCTCTGAGAGAAATCCCTTATTCATGAGGATGTGTGCTAAATAATTC 1608  
 QY 649 LeuLeuAlaArgGlnGlyGly-----AspGlyAsnProPheIleLysLeuSerAsp 666  
 |||||  
 Db 1609 CTGCTTATCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1668  
 QY 667 ProGlyValSerProThrValLeu-----SerLeuGlnMetLeu 679  
 |||||  
 Db 1669 CCTGGCATTTAGCATTTACAGATTCAGAGAGACATTTCTCTCTGTGTTCCAGATTCCT 1728  
 QY 680 ThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeu 699  
 |||||  
 Db 1729 CAGAGAGAAATACCATGGTATCCACCTGAGTGCATTCAGATTCCTAAATAATCTACCTG 1788  
 QY 700 GluAlaAspLysTyrGlnPheGlyAlaThrThrPheLysValPheGlnArgGlyProAla 719  
 |||||  
 Db 1789 GCAACAGACAAAGTGCAGCTTCGGAGCACCTCTGGAGATCTCAGTGGAGAGATAG 1848  
 QY 720 HisIleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGlnAspGlnGlnLeu 739  
 |||||  
 Db 1849 CCCCTGAGTCTCTGATTCCTCAAGAAAGCTGCAATTCATGAAATTAACCACTCCTT 1908  
 QY 740 ProAlaLeuLysTyrThrGlnLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759  
 |||||  
 Db 1909 CCGTACCACCAAGTGCAG 1968  
 QY 760 GlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779  
 |||||  
 Db 1969 GATTTTCAGGCTGCTTCAGAGCTGCATCGATCTTAACACCTGTTACTACCAT 2028  
 QY 780 TyrGlnLeuLeuSerAspProThrProGlyIleProSerProArg---AspGlnLeuCys 798  
 |||||  
 Db 2029 TATGAACTACTACAGAA---AATGACATGCTACCAACATGAGAAATAGTCCCTAGG 2085  
 QY 799 ValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeu 818  
 |||||  
 Db 2086 TTTTCTGTGCTCT-----TTTGAAGACAGGAGACCTTACACAGTTTGAAGAGAGACTTG 2139  
 QY 819 LysTyrIleSerLeuGlnGlyLysGlyAsnPheGlySerValGlnLeuCysArgTyrAsp 838  
 |||||  
 Db 2140 AAGTTTCTACAGAGCTTGGCAAAAGTAACTTCGGAGTGGAGATGTGCGCTATGAC 2199  
 QY 839 ProLeuGlnLysAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPro 858  
 |||||  
 Db 2200 CCCCTGAGAGCAACACTGCGAGGTGTGCTGTGTAAGAAACTCCGACAGACAGCTGAA 2259  
 QY 859 AspGlnGlnArgAspPheGlnArgGlnIleGlnIleLeuLysAlaLeuHisSerAspPhe 878  
 |||||  
 Db 2260 GAGACACCTCCGAACTTTGAGAGGAGATGAGATCTCTGAATCTTTCGACAGCATGAAC 2319  
 QY 879 IleValLysTyrArgLysValSerTyrGlyProGlyArgGlnSerLeuValGluValMet 898  
 |||||  
 Db 2320 ATGCTGAAGTACAGAGAGAGTGTCTACAGTGGGCTGGCGCAACCTTAATTAATATG 2379  
 QY 899 GluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHis 917

Db 2380 GAATATTACCATATGAGAGTTTACAGACATATCTCCAAAACATAAGACGATAGAT 2439  
 QY 918 ThrAspArgLeuLeuPheAlaTyrGlnIleCysLysGlyMetGluTyrLeuGlyAla 937  
 |||||  
 Db 2440 CACAAAACCTTCTCANTACATCATCTCAGATATGCAAGGAGCATGAAATCTTGTGTA 2499  
 QY 938 ArgArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValGluSerGluAlaHis 957  
 |||||  
 Db 2500 AAAAGTATATCCACAGGAGACTGGCAACAGAAACATATGTGTGGAATAAGACAGAGC 2559  
 QY 958 ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVal 977  
 |||||  
 Db 2560 GTTAAATAGAGACTTCGATTAACCAAGCTTTCGCCGAGCAAGAAATACTACAAA 2619  
 QY 978 ValArgGluProGlnGlnSerProIlePheTyrPyrAlaProGluSerLeuSerAspAsn 997  
 |||||  
 Db 2620 GTAAAGAGCCAGGAGGAGAGCCCATATTTCTGTACGACCTGAAATCTTACGAGAGC 2679  
 QY 998 IlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheThr 1017  
 |||||  
 Db 2680 AAGTTTCTGTGGCTCGATGTGTGAGCTTGTGAGTGTCTATACGAACTTTTCA 2739  
 QY 1018 TyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetLeuGlyProGluArg 1037  
 |||||  
 Db 2740 TACATCGAGAGAGTAAAGTCCACCCGCGTGAATTTATGCAAGATTTGCAATGATAA 2799  
 QY 1038 GluGluProProLeuCys---ArgLeuLeuGlnLeuLeuAlaGlnGluArgArgLeuPro 1056  
 |||||  
 Db 2800 CAAGGCAAAATGATGTGTCTTCATTTGATAGAGCTACTAGAGACCAAGCAAGATTCGA 2859  
 QY 1057 ProProProThrCysProThrGluValGlnGluLeuMetGlnCysTyrPheProGlu 1076  
 |||||  
 Db 2860 AGGCCAGAGAGATCCAGATGATGATTTATGATCATGACAGAGTGTCTGCAACAACAT 2919  
 QY 1077 ProHisAspArgProAlaPheAlaThrLeuSer 1087  
 |||||  
 Db 2920 GTGAGCCAGCGTCCCTCTTCAGGAGCTTTC 2952

RESULT 12  
 US-08-446-0108-2  
 ; Sequence 2, Application US/084460108  
 ; Patent No. 5716818  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;  
 ; APPLICANT: Harpur, Ailsa  
 ; TITLE OF INVENTION: No. 5716818el Protein Tyrosine Kinase  
 ; NUMBER OF SEQUENCES: 25  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felle & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York City  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/446,010B  
 ; FILING DATE: 19-May-1995  
 ; CLASSIFICATION: 433  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/446,038  
 ; FILING DATE: 19-May-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/064,067  
 ; FILING DATE: 30-Jun-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US91/08889  
 ; FILING DATE: 26-No. 5716818-1991

```

PRIOR APPLICATION DATA:
: APPLICATION NUMBER: Australian PK3594/90
: FILING DATE: 28-No. 5716818-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: Australian 88229/91
: FILING DATE: 27-No. 5716818-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Baer, Madeline F.
: REGISTRATION NUMBER: 36,437
: REFERENCE/DOCKET NUMBER: LUD 5244.3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-688-9200
: TELEFAX: 212-838-3884
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3495 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: nucleic acid
: US-08-446-010B-2

Alignment Scores:
Pred. No.: 9,43e-210 Length: 3495
Score: 2360.00 Matches: 489
Percent Similarity: 65.89% Conservative: 164
Best Local Similarity: 49.34% Mismatches: 300
Query Match: 40.27% Indels: 38
DB: 1 Gaps: 15

US-09-397-967-16 (1-1099) x US-08-446-010B-2 (1-3495)
OY 128 lleuaspheulhisvalleuglnhisleuphealaglhisargserleuvalser 147
DB 1 ctccttgatgactttgtctatgtcttacccttccscacgtggcgacatgattttttcac 60
OY 148 glyargleuprovalglyleusermetlysgluginglyglupheleuserleuval 167
DB 61 ggatgatataaagactacgtctgacatgacactcagaaagactgctggatggcggtg 120
OY 168 leuaspheulhisleumetalargsluglnalaglnargproglglyleuleuysrhr 187
DB 121 tttagacatgatgacatgacatgacaaagacacagactccacgtcgctgataactgt 180
OY 188 valserlyrlyvalacyleuproproserleuargvalilleglglylnasrhr 207
DB 181 gtccactacacagacattttaccnaagtcgctcgaggaagatccagactatcacatt 240
OY 208 valthrargarglileargatgthrvalalleualaleuleuprocglyargleu 227
DB 241 tttaacccggaagcgaatcagatgacagatttgcagacttcatccaaattcagatgt 300
OY 228 proglargproglargleumetalalysrlylleuasrleuglnargleuvalpro 247
DB 301 aaagcactgccaggaacacttaaaacttaagatgctttataaaccstgaaacccstgacgt 360
OY 248 alalaatthrthrleuvalpheargval-----glyleuproglglyalagln 262
DB 361 gcccttctacacagaaacatttgaagtaaaagaaatctgcagagaggt---cccttcagtgag 417
OY 263 glugluargleuleuvalaglyaspasnlylleprotprserse----- 280
DB 418 gagatgtttgcacacattatgataactggaacggtgaaattcagtgctcagaagsgaaa 477
OY 281 -----asnaspdluleupheglnthrphesysasrpherpoglu 293
DB 478 cataagaaatgacacactgcacagaaacgactacagattatgatttgcattccctgat 537
OY 294 lleualsryvalserlleasnlyalaproargvalglyproalaglyglnhisargleu 313
DB 538 attattgtatgtcagatatttaacacagaaacag---gaatgcctacactgaaagtagagtt 594
OY 314 valthrvalthrargmetleupglynhisleuglnalagluclupheproglyleuproglu 333

595 gtgaccgtccacagacagcggaaagcttggaaataagaaacttaccatttaaaagaa 654
OY 334 alaleuserphevalalaleuvalaspglyrphargleuilecysasrserarghis 353
DB 655 gccctgtcatgtcgtgcatattatgacgggtatgacagactaaacgacgacacacat 714
OY 354 tyrphecyslysgluvalalaproproargleuleugluginalalaspvalcysnhis 373
DB 715 tacctctgcgaagagagtgctcccccactgtgctgcgaacacatcacagacactgcac 774
OY 374 glyproillethrleuasrphelalallehislyseuylsvalalaglyserleuprogly 393
DB 775 gcccaattttcattgatttggccatcagcaaaactaaagaaagcagaaacacacacgga 834
OY 394 thrtyrilleuargargserproglinsprlyrasrpheluleuthralcysval 413
DB 835 ctgtatgtacttgcattgtagcccttaagacattcaacaaatgcttcgaccttggccgtt 894
OY 414 glnthrproleuglyproasptryllysglycysleuileargglinsrproserglyala 433
DB 895 gagcagaaatgattatgaaatataaacactgttgcattacaaagaaatgaaatgagacag 954
OY 434 pheuserleuvalglyleuserglinprohisargserleuargleuleuvalalacys 453
DB 955 tacacactcagatggactaacagaaacttcagtagtctttaagacacttttgaaattgctac 1014
OY 454 trpanserglyleuargvalaspglyalalaleuvalyleuthrsercysglsalpro 473
DB 1015 cagatggaaactgtgcctcagacagatgacttccacagttccacaaatgctctcccca 1074
OY 474 argprolysglylyserasnleuilevalargarglycysasnproalaproala 493
DB 1075 aagcgaagaaataatgaaacacttctgtctcaga-----acaaatgctgttgcatt 1128
OY 494 procllysserprosercyscysala-----leuthrlnleuserphenisrhr 510
DB 1129 gttcagctctccacacacatttacagagcattataatgataaaatggttttcacnaa 1188
OY 511 lleprothrpsrserleugluthrhisgluasleuclynhislyserphenrlysrile 530
DB 1189 atcaggaatgaaatgatttgaatttaagaaaccccttgccacagacacttttacaaaaata 1248
OY 531 phearglyserargargvalvalasp---glgluthrhisasrsergluvalleu 549
DB 1249 tttaagcttaagaaacagaaagttgagattatgctcagctgcacgaaacccgaagttctt 1308
OY 550 leuulysvalmetasrserarghisargasnlycysmetcgluserpheleucluaalasr 569
DB 1309 ttgaagctcctagatgaagacacatgaamaactatttcagagctttcttgaaacagacag 1368
OY 570 leuuserleuvalserlyrprohisleuvalleuleuhsiglyvalcysmetalagly 589
DB 1369 atcatgtagtcagctttctcacaagcatttgggtttgaaatgaaatgaaatgtagtctgtga 1428
OY 590 ---aspserlleuvalglinluphevaltyrleuglyalalleasrphenrlyleuarg 608
DB 1429 gagcagaacattttgcttcagaagctttgaaatttggatctcagctgatacattcagag 1488
OY 609 lysargglynhisleuvalserlasertrplysleuclnvalthrlysglnleuvalatyr 628
DB 1489 aagaacaaaatgcttaataatgattatgtaaaacttgcagtgccgaacagattggcattgg 1548
OY 629 alaleuasntyrleuglnasrlysglyleuprohisgllyasnvalseralarglyval 648
DB 1549 gccatgcactctctcgaagaaatcccttatttcattgggaatgctgtgctgtaaaatgattgc 1608
OY 649 leuleuvalargglycylgyl-----aspglyasnproprophenileuysleuserarg 666
DB 1609 ctgccttatgcagaaagaaacagacagaaacgggaacccacacttgcatacnaacttattgat 1668
OY 667 proglyvalserprothrvalleu-----serleuglnmetleu 679
```

Db 1669 CTTGCATAGACATACAGTTCTACCGAAGACATTTCTTCTGTTGTTCCAAAGTTCTT 1728  
 Oy 680 ThAspArgIleProtrpValAlaIleProGluCysLeuGlnLualGlnThrLeuCysLeu 699  
 Db 1729 CAGAGGAATACCATGTGACCACTGAGTGTGATGAGATCTTAATAATCTAACTACTG 1788  
 Oy 700 GluAlaAspIlystrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgIlyProAla 719  
 Db 1789 GCACACACACAGAGTGCCTCGGAGCCTCTGAGGAGATCTGCAGTGGAGAGATTAAG 1848  
 Oy 720 HistIleThrSerLeuGluProAlaIleLeuArgPheLeuGlnPheTrpGluAspGlnGlnLeu 739  
 Db 1849 CCCCTGAGTGTCTGATCTCAAGAAAGCTGAGTTCTTATGAGATTAAGCATCACTT 1908  
 Oy 740 ProAlaLeuIlystrpTrpGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyraAspPro 759  
 Db 1909 CCTGCACCCAGATGGACAGATTTGCAAACTTAAATTAATTCAGTGCAGCTATGAGCCA 1968  
 Oy 760 GlyArgArgProSerPheArgAlaIleLeuArgPheLeuGlnGlnLeuIleThrSerAsp 779  
 Db 1969 GATTCAGGCTGCTTTCAGAGCTGTACCGTATCTTAACCTGTTTACTTCCAGAT 2028  
 Oy 780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCys 798  
 Db 2029 TATGAACTACTTAACAGAA---AATGACATGCTACCAAAACATGAGATAGTGCCTTAGG 2085  
 Oy 799 ValAlaGlyAlaGlnLeuTyraCysGlnAspProAlaIlePheGluGlnArgIleLeu 818  
 Db 2086 TTTTCTGTGCTGCT---TTTGAAACAGAGGAGCCTTACACAGTTTGAAGAGACACTTG 2139  
 Oy 819 LysTrpIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyraAsp 838  
 Db 2140 AAGTTTCTACACACACTTGGCAAGTAACTTCCGGAGTGTGGAGATGTGCCGCTATGAC 2199  
 Oy 839 ProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPro 858  
 Db 2200 CCGGTGACAGAGACACACTGTGCGAGCTGCTGCGTGAAGAACTCCACACAGCACTGAA 2259  
 Oy 859 AspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPhe 878  
 Db 2260 GAGCAGCTCCGAGACTTGTAGAGGAGATCGAGATCTTGAATCTTTCAGCATGACACAC 2319  
 Oy 879 IleValLysTyraGlyValSerTyraGlyProGlyArgGlnSerLeuArgLeuValMet 898  
 Db 2320 ATCGTAACTACAGGAGATGTGTACAGTGTGCGGCTGCGCAGCACTAAGATTAATATG 2379  
 Oy 899 GluTyraLeuProSerGlyCysLeuArgAspLeuGlnArgIleHisArg---GlyLeuHis 917  
 Db 2380 GAATATTACCATATGAGATTGAGACTATCTCCAAACATTAAGAAAGCATAGAT 2439  
 Oy 918 ThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyraLeuGlyAla 937  
 Db 2440 CACAAAACCTTCTCAATACACATCTCAGATATGCCAAGGATGAAATATCTTGATACA 2499  
 Oy 938 ArgArgCysValHisArgAspLeuAlaIleArgAsnIleLeuValGlnSerGluAlaHis 957  
 Db 2500 AAAAGGTATATCCACAGGAGCTGCGACACAGAGAAATATGTTGGAATAATGANAACAG 2559  
 Oy 958 ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyraVal 977  
 Db 2560 GTTAATAATAGAGACTTCGATTACCAAAAGCTTCCCGAGAGACAAGAAATACTACAA 2619  
 Oy 978 ValArgGluProGlyLysSerProIlePheTrpTyraAlaProGlnSerLeuSerAspAsn 997  
 Db 2620 GTAAGAGAGCCAGGAGAAACCCCATATCTGTTAGCACACCTTAATCTTGAGGAGACC 2679  
 Oy 998 IlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyraGluLeuPheThr 1017  
 Db 2680 AAGTTTCTGTGCGCTCAGATGTGTGAGCTTTGAGTGTCTTATACGACTTTTCACA 2739  
 Oy 1018 TyrcysAspIlySerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArg 1037  
 Db 2740 TACATCGAAGAAAGCTAAAGTCCACCGCTGGAATTTATGCGAATGATGCAATGATAA 2799

Oy 1038 GluGlyProProLeuCys---ArgLeuLeuGlnLeuAlaGlnGlyArgArgLeuPro 1056  
 Db 2800 CAAGGCAAAATGATGTGTTCCATTGTATAGAGTACTAGAGACCAAGCAAGATTTGCCA 2859  
 Oy 1057 ProProProThrCysProThrGluValGlnGlnLeuMetGlnLeuCysTrpAlaProGlu 1076  
 Db 2860 AGGCCAAGAGATGCCAGATGATGATTAATGATCATGATGATGATGATGATGATGATGAT 2919  
 Oy 1077 ProHisAspArgProAlaPheAlaThrLeuSer 1087  
 Db 2920 GTGAGCCAGCTGCTCTCTTCCAGGACCTTTC 2952  
 RESULT 13  
 US-08-805-445-2  
 Sequence 2, Application us/08805445  
 Patent No. 5821069  
 GENERAL INFORMATION:  
 APPLICANT: Walks, Andrew F.; Ziemlecki, Andrew;  
 APPLICANT: Harput, Alisa  
 TITLE OF INVENTION: No. 5821069e1 Protein Tyrosine Kinase  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Felfe & Lynch  
 STREET: 805 Third Avenue  
 CITY: New York City  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10022  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
 COMPUTER: IBM PS/2  
 OPERATING SYSTEM: PC-DOS  
 SOFTWARE: WordPerfect  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/805,445  
 FILING DATE: 25-FEB-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/446,038  
 FILING DATE: 19-MAY-1995  
 APPLICATION NUMBER: 08/064,067  
 FILING DATE: 30-Jun-1993  
 APPLICATION NUMBER: PCT/US91/08889  
 FILING DATE: 26-No. 5821069-1991  
 PRIOR APPLICATION DATA: Australian PK3594/90  
 APPLICATION NUMBER: 5821069-1990  
 FILING DATE: 28-No. 5821069-1991  
 PRIOR APPLICATION DATA: Australian 88229/91  
 APPLICATION NUMBER: 5821069-1991  
 FILING DATE: 27-No. 5821069-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Hanson, No. 5821069man D.  
 REGISTRATION NUMBER: 30,946  
 REFERENCE/DOCKET NUMBER: LUD 5244  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-688-9200  
 TELEFAX: 212-838-3884  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3495 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: nucleic acid  
 US-08-805-445-2  
 Alignment Scores:  
 Pred. No.: 9,43e-210 Length: 3495  
 Score: 2360.00 Matches: 489  
 Percent Similarity: 65.89% Conservative: 164  
 Best Local Similarity: 49.34% Mismatches: 300  
 Query Match: 40.27% Indels: 38











QY	740	ProAlaLeuGlySerPheTrpGluLeuAlaGlyLeuLeuIleHnGlnCysMetAlaIleValysPro	759
Db	1909	CGTGACCCCAAGTGGACAGAGATTGGCAAACTTATTAATATGTGCATGTGACATATGACCA	1968
QY	760	GlyArgArgProSerPheIleArgAlaIleLeuAlaArgAspLeuAlaGlyLeuLeuIleHnSerAsp	779
Db	1969	GATTTCACGGCTCGCTTTCACAGCTGTGATCCGGATCTTAAACGCGCTGTTTACTCCAGAT	2028
QY	780	TyrGluLeuLeuSerAspProTrpProGlyIleProSerProArg---AspGluLeuCys	798
Db	2029	TATGAACTACTAACAAGAA---AATGACATGCTACCAAAACATGAGAAATAGTGCCTTAAGG	2085
QY	799	ValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluAlaArgHnIleu	818
Db	2086	TTTTCTGGTGTCT	2139
QY	819	LysTyrIleSerLeuLeuGlnGlyLysGlyAsnPheGlySerValGluLeuLeuSerTyrTyrAsp	838
Db	2140	AAGTTTCTACAGCAGCGTTGGCAAAAGTAACCTTGGGAGGTGTGAGATGTGCCGCTATGAC	2199
QY	839	ProLeuGlyAspAsnHnTrpGlyProLeuValAlaValLysGlnLeuGlnHnIleSerValPro	858
Db	2200	CCGGTCAGGAGCAACACTGGCGAGGTGTGCTGTGATGAAGAACTCCAGCAGCAGCTGAA	2259
QY	859	AspGlnAlaArgAspPheGlnIleArgGluIleGlnIleLeuLysAlaLeuHnIleSerAspPhe	878
Db	2260	GAGCACCTCCGAGACTTTGAGAGGGAGATGAGATCTGTAATCTTGGCAGATACAAAC	2319
QY	879	IleValLysTyrArgGlyValSerTyrGlyProGlyIleArgGlnSerLeuArgLeuValMet	898
Db	2320	ATCGCAAGTACAGAGGAGAGTGTCTACAGTGGCGGGCGCGCAACTGAACTTAATATATG	2379
QY	899	GluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHnIleArg---GlyLeuHnIle	917
Db	2380	GAATTTTATCCATATGTGAAGTTTACGAGACTATCTCCAAAACATTAACAAACGCGTAAT	2439
QY	918	ThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAla	937
Db	2440	CACAAAACCTTTTCAATACATCATCTCAGATATGCAAGGGCATGTGAATATCTTGTTACA	2499
QY	938	ArgArgCysValIleHisArgAspLeuAlaAlaIleArgAsnIleLeuValGluSerGluAlaHnIle	957
Db	2500	AAAAGTATATCCACAGGAGCTGGCAACAGAGAACATATTGGTGGAAAATGAGAAACAG	2559
QY	958	ValLysIleIleAspPheGlyLeuAlaIleLeuLeuProLeuGlyLysAspTyrTyrVal	977
Db	2560	GTTAAATAGAGACACTTCGGATTACCAAAAGCTTGTCCCGCAGGACCAAGAAATACTACAA	2619
QY	978	ValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsn	997
Db	2620	GTAAGGAGCCAGGAGGAAACCCCATATTCTGTACGCACTGAATCCCTTTCAGGAGAG	2679
QY	998	IlePheSerArgGlnIleAspValIleTrpSerPheGlyValValLeuTyrGluLeuPheHnIle	1017
Db	2680	AAGTTTCTGTGGCCTCAGATGTGTGGAGCTTGGAGTGGTCTTATACCAACTTTTACA	2739
QY	1018	TyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetCylProGluArg	1037
Db	2740	TACATCGAGAGAGTAAAGATCCACCCGTCGATATTATGCGAATGATTTGGCAATGATPAA	2799
QY	1038	GluGlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluLysArgArgLeuPro	1055
Db	2800	CAGAGCCAAATGATGTGTTCCATTTGATATGACCTTACTAAGCAACGAAGATTTTCCA	2859
QY	1057	ProProProTrpCysProTrpGluValaGlnGluLeuMetGluLeuCysTrpAlaProGlu	1076
Db	2860	AGGCCAAGAAGATCCCAAGATGAGATTATGATCATGACAGAGTGTGCGAACAACAAT	2919
QY	1077	ProHisAspArgProAlaPheAlaTrpLeuSer	1087
Db	2920	GTGAGCCAGCGCTCCCTCTTCAGGAGCACTTTC	2952

Tue Apr 29 06:04:28 2003

us-09-397-967-16.rni

Page 30

Search completed: April 28, 2003, 17:00:17  
Job time : 296.669 secs

---

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 28, 2003, 18:32:45 ; Search time 3018 Seconds

(without alignments)  
5897.561 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 1099

Sequence: 1 MAPSEPTPLIPORCSLSS.....RPAFATISPOLDPLMRGPG 1099

Scoring table:

OLIGO	
Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 1

Total number of hits satisfying chosen parameters: 32308016

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 150 summaries

Command line parameters:

```

-MODEL=frame_plus_p2n.model -DEV=xlp
-O=/gen2.1/USPRO.spool/US09397967/runat_26042003.182927.11215/app.query.fasta.1.1287
-DB=FAST -QEXT=fastext -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=150
-DOCALLIGN=200 -THR=SCORE=quality -THR_MIN=1 -ALIGN=100 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09397967 @CGN.1.1.2605 @runat_26042003.182927.11215 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=10 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

```

Database :

```

ESR:*
1: em_estba:*
2: em_esthum:*
3: em_estlm:*
4: em_estlmv:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	13.8	971	13	BI411962 602966375
2	149	13.6	580	10	BB608771 BB608771
3	142	12.9	583	10	BB631706 BB631706
4	142	12.9	720	12	BG873355 BG873355
5	135	12.3	1625	11	BC027234 BC027234
6	110	10.0	512	9	AA510093 AA510093
7	92	8.4	545	12	BE916148 BE916148
8	88	8.0	826	13	BI57690 BI57690
9	82	7.5	675	12	BF150250 BF150250
10	77	7.0	907	12	BG870513 BG870513
11	74	6.7	453	9	AA023709 AA023709
12	74	6.7	478	9	AA023670 AA023670
13	73	6.6	593	9	AA881653 AA881653
14	71	6.5	698	14	BM935182 BM935182
15	69	6.3	436	12	BF181053 BF181053
16	65	5.9	436	10	AW141115 AW141115
17	65	5.9	590	14	W48204 W48204
18	62	5.6	559	12	BG800853 BG800853
19	60	5.5	358	13	BI153609 BI153609
20	56	5.1	621	10	AW742526 AW742526
21	53	4.8	539	12	BE913247 BE913247
22	51	4.6	867	9	AA755769 AA755769
23	49	4.5	543	12	BE753801 BE753801
24	49	4.5	557	13	BI339676 BI339676
25	46	4.2	544	12	BF705690 BF705690
26	46	4.2	562	12	BF192178 BF192178
27	45	4.1	659	17	AZ958543 AZ958543
28	44	4.0	417	13	BI540725 BI540725
29	44	4.0	441	10	AA020387 AA020387
30	44	4.0	501	10	BE623895 BE623895
31	44	4.0	597	9	AA015273 AA015273
32	42	3.8	581	12	BF149581 BF149581
33	41	3.7	554	12	BF191625 BF191625
34	39	3.5	731	9	AA920299 AA920299
35	38	3.5	410	9	AA023069 AA023069
36	35	3.2	308	9	AA174246 AA174246
37	34	3.1	231	10	AW501263 AW501263
38	34	3.1	254	10	AW501582 AW501582
39	34	3.1	286	10	AW502902 AW502902
40	34	3.1	437	10	AW501452 AW501452
41	34	3.1	451	10	AW500984 AW500984
42	34	3.1	462	10	AW408832 AW408832
43	31	2.8	478	12	BF546077 BF546077
44	30	2.7	483	10	AW462965 AW462965
45	30	2.7	763	9	AJ399251 AJ399251
46	30	2.7	769	9	AJ442008 AJ442008
47	28	2.5	411	10	AW503697 AW503697
48	27	2.5	491	12	BF441858 BF441858
49	27	2.5	550	10	BE309453 BE309453
50	25	2.3	526	12	BF191620 BF191620
51	25	2.3	564	13	BM088287 BM088287
52	25	2.3	671	14	BO709825 BO709825
53	24	2.2	371	10	BB791980 BB791980
54	23	2.1	339	10	AW230832 AW230832
55	22	2.0	540	10	AW505513 AW505513
56	21	1.9	472	10	AW785501 AW785501
57	18	1.6	258	10	BB590115 BB590115
58	18	1.6	291	9	A1651109 A1651109
59	18	1.6	319	13	BM032194 BM032194
60	18	1.6	373	13	BM146700 BM146700
61	18	1.6	413	10	AW504672 AW504672
62	18	1.6	481	17	AO192794 AO192794
63	18	1.6	781	13	BI906136 BI906136
64	18	1.6	814	13	BI911448 BI911448

```

C 65 17 1.5 313 12 BF147824 uY83609..x
C 66 17 1.5 426 13 BG995133 CM0-HT129
C 67 17 1.5 449 9 A1472755 tA13e03..x
C 68 17 1.5 449 9 A1596913 EST250616
C 69 16 1.5 239 10 BB285471 BB285471
C 70 16 1.5 258 12 BF367544 IL5-GN003
C 71 16 1.5 304 10 BB285470 BB285470
C 72 16 1.5 398 9 AA453345 zxA4h10..x
C 73 16 1.5 406 14 T48592 Phb6_19/1T
C 74 16 1.5 430 13 BM147781 TCANP1014
C 75 16 1.5 532 12 BG364706 GC575a07..y
C 76 16 1.5 612 9 AA521703 v107907..x
C 77 16 1.5 753 12 BG870416 602791325
C 78 16 1.5 766 13 B1153342 B0899824 AGENCOURT
C 79 16 1.5 920 14 B0899824 AGENCOURT
C 80 16 1.5 945 14 BQ732700 AL272087 Tetradon
C 81 16 1.5 989 17 CNS038Mw AL232817 Tetradon
C 82 16 1.5 1001 17 CNS038Mw BF373713 PM4-SN001
C 83 15 1.4 253 12 BF373713 PM4-SN001
C 84 15 1.4 754 17 BH043390 RPCI-24-3
C 85 14 1.3 281 10 AM437524 78773 MAR
C 86 14 1.3 296 10 AW212845 uo68402..x
C 87 14 1.3 318 9 AA008555 mg6b10..x
C 88 14 1.3 349 10 BB711016 BB711016
C 89 14 1.3 636 14 BQ780982 UT-R-FEO-
C 90 14 1.3 683 14 BQ781477 UT-R-FEO-
C 91 14 1.3 881 13 B1643881 NXPV_126-
C 92 13 1.2 441 12 BG509232 sac87h10
C 93 13 1.2 444 12 BM813523 EST591616
C 94 13 1.2 500 13 B1424599 sah53e02..
C 95 13 1.2 562 10 AM231037 uo70a12..y
C 96 13 1.2 145 10 BE156215 OVO-HT036
C 97 12 1.1 214 9 AA320481 AA320481
C 98 12 1.1 270 14 H24996 y139g05..r1
C 99 12 1.1 294 13 B1051048 CM2-GN028
C 100 12 1.1 308 10 BB605196 BB605196
C 101 12 1.1 316 10 BE554456 ur42910..y
C 102 12 1.1 318 12 BF741156 CM1-HB000
C 103 12 1.1 324 14 D77965 MUSC0C03 mo
C 104 12 1.1 336 9 AA069054 zmi0h02..x
C 105 12 1.1 346 12 BF886669 RC6-TN007
C 106 12 1.1 353 9 AA715499 nv73c09..x
C 107 12 1.1 354 17 BH826047 BACP23-E
C 108 12 1.1 354 10 AM059494 fe14407..y
C 109 12 1.1 366 9 AA233516 zr29h06..x
C 110 12 1.1 372 9 AA572234 v151905..x
C 111 12 1.1 383 9 AL138105 AL138105
C 112 12 1.1 392 9 A1483943 EST249814
C 113 12 1.1 399 14 R17980 y916f02..r1
C 114 12 1.1 409 17 CNS04002 AL300284 Tetradon
C 115 12 1.1 437 14 BM823021 K-EST0093
C 116 12 1.1 437 14 BM857398 K-EST0141
C 117 12 1.1 438 10 AV978849 AV978849
C 118 12 1.1 445 13 BG987835 CM3-HT115
C 119 12 1.1 453 12 BF391056 UT-R-CAL-
C 120 12 1.1 462 13 B1221056 602937661
C 121 12 1.1 472 13 BM288735 530131 MA
C 122 12 1.1 472 13 BF774038 283572 MA
C 123 12 1.1 473 12 BG458915 947055E04
C 124 12 1.1 475 14 BM968852 UT-H-COO-
C 125 12 1.1 479 12 BF769324 RC3-TT001
C 126 12 1.1 488 9 A1444260 f4a0f06..y
C 127 12 1.1 496 10 AM665954 h110c10..x
C 128 12 1.1 497 14 B0560624 HA065D12..x
C 129 12 1.1 498 10 AM469824 hb35f12..x
C 130 12 1.1 499 17 A2499369 lW0337M03
C 131 12 1.1 502 10 AM632498 92470 MAR
C 132 12 1.1 504 13 BJ478073 BJ478073
C 133 12 1.1 504 14 R56713 y995a06..r1
C 134 12 1.1 517 9 A1150337 qf35b04..x
C 135 12 1.1 517 10 AM938783 PM3-DT006
C 136 12 1.1 518 14 BM756676 K-EST0035
C 137 12 1.1 520 12 BE872391 601448652

```

```

138 12 1.1 523 10 AM489577 UI-M-BH3-
C 139 12 1.1 526 12 BF736600 OV3-KT001
C 140 12 1.1 528 17 CNS01XHO AL171693 Tetradon
C 141 12 1.1 532 12 BG798512 BG798512
C 142 12 1.1 532 17 AO377376 RPCI-11-1
C 143 12 1.1 536 14 BM756648 K-EST0035
C 144 12 1.1 538 12 BG812095 BG812095
C 145 12 1.1 543 10 BE330992 dab99c01..
C 146 12 1.1 546 12 BE330992 soc92c05..y
C 147 12 1.1 546 12 BE722796 191178 MA
C 148 12 1.1 547 14 BM742052 BM742052
C 149 12 1.1 548 14 BQ399048 BQ399048
C 150 12 1.1 548 14 BM756028 BM756028
C 150 12 1.1 549 10 BE632486 uv57c05..y

ALIGNMENTS
RESULT 1
BI411962 971 bp mRNA linear EST 14-AUG-2001
LOCUS 602966375F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5121786 5',
DEFINITION mRNA sequence.
ACCESSION BI411962
VERSION BI411962.1 GI:15172885
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 971)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gcaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M1297 row: e column: 19
High quality sequence start: 25
High quality sequence stop: 785.
Location/Qualifiers
1. 971
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5121786"
/clone_lib="NCI CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; site_1: NotI; site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCCGCTGTCTTTTCTTTTCTTTT 3']
Double-stranded cDNA was ligated to Eco RI-adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 216 a 274 c 279 g 202 t
ORIGIN
Alignment Scores: 1.19e-137 length: 971
Score: 152.00 Matches: 196
Percent Similarity: 98.99% Conservative: 0
Best Local Similarity: 98.99% Mismatches: 1

```

Query Match:	13.83%	Indels:	2
DB:	13	Gaps:	0
US-09-397-967-16 (1-1099) x B1411962 (1-971)			
QY 490	ProAlaProAlaProGlyCysSerProSerCysCysAlaLeuThgInleuSerPheHis	509	
Db 51	CCCCGGCGTGGCCCGGGGCTGCTCCCGCTCTCTGCGCTACACAGGTAGCTTCCAC	110	
QY 510	ThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGlnHisGlySerPheThrIys	529	
Db 111	ACAAATTCACAGCGACAGCGCTGGATGGGACAGCAACCTGGGTCACAGGTCTTTTCCAAAG	170	
QY 530	IlePheAerGlySer-AerArgGluValValaspGlyGluTrpHisAspSerGluValle	549	
Db 171	ATCTTCGGTGGCCG-CAGCGGGAGGTGCTGGATGCTGAGACACATGACTCGGAAGTCTT	229	
QY 549	UleuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSe	569	
Db 230	CCTGAAGGTCATGAGCTCTCAGACATCGGAAACGTGAGTCTTTTTCGGAAGCCGCAAG	289	
QY 569	FluMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaG1	589	
Db 230	CTTGATGAGCCACAGATGATCTTACCCTGGCTGTACTGCGCGGCGTGTGCATGGCTGG	349	
QY 589	YaspSerIleMetValGlnGluPheValTyrLeuGlyAlaAlaAspMetTyrLeuArgIly	609	
Db 350	AGACAGCATCATGTGGCGAGGAATTTGTGTATCTAGGAGCAATTGACATGTACTCGGCA	409	
QY 609	SargGlyHisLeuValSerAlaSerTrpIysLeuGluValTrpIysGlnLeuAlaTyrAl	629	
Db 410	GGGTGGCCACCTGGGGTGTACGCCAGCTGGAAATCGAGGAGACCAAGCAGCTGGCATATGC	469	
QY 629	AlenAsnTyrLeuGlnAspIysGlyLeuProHisGlnLysAsnValSerAlaArgLysValle	649	
Db 470	CCTTACTACTTGGAGGACAAAGGCTTCTCTCAGCGCAACGTCTCAGCAGGAAAGTGTCT	529	
QY 649	UleuAlaArgGluGlyGlyAspGlyAsnProProPheIleIleLysLeuSerAspProGlyIya	669	
Db 530	CTGTGCTGTGTGGGGGGGTGATGGGAATTCACCTTTCATTAAAGCGAGGATCCGTGTGT	589	
QY 669	IserProThrValLeuSerLeuGluMetLeuThrAspArgIleProTrpVal 686		
Db 590	CAGTCCCACTGTGCGTGGACCTCGGAATCTCTCCAGCACGAATACCTGGGTG 641		
RESULT 2			
BB608771	580 bp mRNA linear EST 26-OCT-2001		
LOCUS	BB608771 RIKEN full-length enriched, 2 days pregnant adult female		
DEFINITION	ovary Mus musculus cDNA clone E330007L03 5', mRNA sequence.		
ACCESSION	BB608771		
VERSION	BB608771.2 GI:16451108		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	1 (bases 1 to 580)		
	Arakawa,T., Carrinced,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,		
	Himoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,		
	,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,		
	Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,		
	Tadami,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,		
	Tegami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,		
	Muramatsu,M. and Hayashizaki,Y.		
	RIKEN MOUSE ESTs (Arakawa,T., et al. 2001)		
TITLE	Unpublished (2001)		
JOURNAL	On Dec 6, 2000 this sequence version replaced gi:11563948.		
COMMENT			

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
url: http://genome.gsc.riken.go.jp/  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. *Genome Res.* 10 (10), 1617-1630 (2000)  
Wagci, K., Fujiwake, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E.,  
Watahiki, M., Itoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*  
10 (11), 1757-1771 (2000)  
Komno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,  
Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)  
Kondo, S., Shitagawa, A., Saito, T., Kiyosawa, H., Yamana, I., Aizawa  
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
Hayashizaki, Y.  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences *Mamm. Genome*, 12, 673-677 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Division of Experimental Animal Research Laboratory in RIKEN.  
Prepare mouse tissues.

```

FEATURES
Source
    Location/Qualifiers
      1..580
        /organism="Mus musculus"
        /dp_xref="taxon:10090"
        /clone="E330007L03"
        /clone_lib="RIKEN full-length enriched, 2 days pregnant
adult female ovary"
        /sex="female"
        /tissue_type="ovary"
        /dev_stage="2 days pregnant adult"
        /lab_host="DH10B"
        /note="Site_1: SalI. Site_2: BamHI. cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5']
GAGGAGAGAGACGGCCGACACTCGACGTTTTTTTTTTTTTTTAA 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adaptor of sequence [5']
GAGGAGAGAGATCTCGACGTTAAATTAATTAATGCCCCCCCCCCC 3'}. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT      100 a      171 c      156 g      153 t
ORIGIN
Alignment Scores:
  Pred. No.:      5 59e-135      Length:      580
Score:           149.00      Matches:      149
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      13.56%      Indels:      0
DB:               10      Gaps:      0
US-09-397-967-16 (1-1099) x BB608771 (1-580)
y      1 MetaIaProPseRgiUgIuThrProLeuIleProGlnArgSerCysSerLeuSerSer 20

```





```

OY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrcysHisArgPheGlyLeu 120
    |||||||
DB 322 AGGCTACGGCTTTATTTCCCTGACTGCTGTGGGCTGAGACATGTCACCGCTTTGGGCTG 361
OY 121 ArgGlyAspLeuThrSerAlaIleuAspLeuHisValLeuGluHisLeuPheAlaGln 140
    |||||||
DB 382 CGCAAGATTTGACACAGTGCATCTTGACTACAGCTTTTAGACATGCTCTTGGCTCAG 441
OY 141 HisArg 142
    |||||||
DB 442 CACCGC 447

RESULT 4
BG873355 720 bp mRNA linear EST 29-MAY-2001
LOCUS BG873355
DEFINITION BG02794328F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4925598 5',
    mRNA sequence.
ACCESSION BG873355
VERSION BG873355.1 GI:14223895
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
    1 (bases 1 to 720)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM10850 row: 9 column: 07
High quality sequence stop: 709.
FEATURES
    source location/Qualifiers
        1..720
            /organism="Mus musculus"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:4925598"
            /clone_lib="NCI_CGAP_SG2"
            /lab_host="DH10B (T1 phage-resistant)"
            /note="Organ: salivary gland; Vector: pCMV-SPOrt6; Site:1;
            Note: Site:2: Sal; Cloned unidirectionally. Primer: Oligo
            dt. Average insert size 1.3 kb. Constructed by Life
            Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 123 a 239 c 229 g 129 t
ORIGIN
Alignment Scores:
Pred. No.: 5.15e-128 Length: 720
Score: 142.00 Matches: 142
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.92% Indels: 0
DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BG873355 (1-720)
OY 916 LeuHisThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysGlySerGlyMetGluTyrLeu 935
    |||||||
DB 79 CTGCACACCGACCGCTACGCTGCTTCGCTTGCACATCTGCACAGGCGCATGAGTACCTG 138
OY 936 GLYALARGARGCYValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGlu 955
    |||||||
DB 139 GGTGGCGCGCGCTGCTACACCTGACCTGCTGCGCGCACATCTTTGGTGGAGAGCGAG 198

```

```

OY 956 AlaHisValysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGluGlyAspTyr 975
    |||||||
DB 199 GCTCATGTGAAGATCCGGAGCTTCGGCCTCGCTTAAGCTGCTGCCCTGGGAAAGACTAC 258
OY 976 TyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSer 995
    |||||||
DB 259 TACGTGTCCGCGAGCGCTGCGCAAGCCCATCTTTGGATATGCCCGGAGTCCCTATCT 318
OY 996 AspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeu 1015
    |||||||
DB 319 GACACATCTTCTCCGCGCATCTGAGCTTGAGCTTCGAGCTGGTGTGTACGAGCTC 378
OY 1016 PheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuAlaGlyMetGlyPro 1035
    |||||||
DB 379 TTCACCTACGCGACACAAGACTCTGAGCCCATCCGCTGAGTTCCTGCGCATGATGGGCGCT 438
OY 1036 GUAArgGluGlyProProLeuCysArgLeuGluGluLeuAlaGluGlyArgGlu 1055
    |||||||
DB 439 GAGCGTGAAGAGACCCCGCTCGCTCCGCTCTGAGCTGCGCAGAGCGCGACGCTC 498
OY 1056 ProPro 1057
    |||||||
DB 499 CCACCA 504

RESULT 5
LOCUS BC027234 1625 bp mRNA linear HTC 07-AUG-2002
DEFINITION BC027234
ACCESSION BC027234
VERSION BC027234.1 GI:20071020
KEYWORDS HTC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC Project URL: http://mgc.nci.nih.gov
    1 (bases 1 to 1625)
AUTHORS National Institutes of Health, Mammalian
JOURNAL Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgebcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Huliyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Wuzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRK Plate: 35 Row: 9 Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein
This clone has the following problem: Incomplete processing.
FEATURES
    source location/Qualifiers
        1..1625
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /map="C57BL/6J"
            /clone="IMAGE:3489805"
            /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
            old, gross tissue."
            /clone_lib="NCI_CGAP_Mam5"

```



```

RESULT 7
LOCUS BE916148 545 bp mRNA linear EST-29-SEP-2000
DEFINITION 601666763p1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:396903 5',
ACCESSION BE916148
VERSION BE916148.1 GI:10416498
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 545)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM9140 row: m column: 16
High quality sequence stop: 545.
FEATURES
location/Qualifiers
1..545
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:396903"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site:1; Salt;
Site:2: NOTI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 93 a 176 c 166 g 110 t
ORIGIN
Alignment Scores:
Pred. No.: 3,29e-79 Length: 545
Score: 92.00 Matches: 92
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.37% Indels: 0
Gaps: 0
DB: 12
US-09-397-967-16 (1-1099) x BE916148 (1-545)
QY 374 GlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGly 393
DB 206 GGACCCATCAGCTAGACTTTCATCCACAGCTGAAGCGCGCTGCCCTCCACAGGC 265
QY 394 ThrTyrIleLeuArgSerProGlnAspTyrAspSerPheLeuLeuThalAlaCys 413
DB 266 ACCATATATCTCCCGCCGACGCGAGCACTATGACAGCTTCTTACCCCTGCGCTC 325
QY 414 GlnThrProLeuGlyProAspTyrLysGlyLysLeuIleArgGlnAspProSerGlyAla 433
DB 326 CAGACTCTCTTGGCCCGGACACTACAAAGGGCTCTCTATCCGCCAGAGACCCGCGGGCT 385
QY 434 PheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGlnLeuLeuAlaIaCys 453
DB 386 TTCTCCCGGTGGCTGCTACGACGCCACCAAGCCCTGCGGAGACTGCTTCACAGCTTC 445
QY 454 TrpAsnSerGlyLeuArgValAspGlyAlaAlaLeu 465

```

```

DB 446 TCGAATTCGTGGCTGCGAGTAGACGCTGCTGCCCTG 481
RESULT 8
LOCUS B1557690 826 bp mRNA linear EST-05-SEP-2001
DEFINITION 603237050f1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5289949 5',
ACCESSION B1557690
VERSION B1557690.1 GI:15445004
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 826)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHM11733 row: d column: 14
High quality sequence stop: 760.
FEATURES
location/Qualifiers
1..826
/organism="Mus musculus"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5289949"
/clone_lib="NCI_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SPORT6; Site:1; Salt;
Site:2: NOTI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
BASE COUNT 147 a 252 c 240 g 187 t
ORIGIN
Alignment Scores:
Pred. No.: 4.45e-75 Length: 826
Score: 88.00 Matches: 208
Percent Similarity: 98.58% Conservative: 0
Best Local Similarity: 98.58% Mismatches: 3
Query Match: 8.01% Indels: 3
Gaps: 0
DB: 13
US-09-397-967-16 (1-1099) x B1557690 (1-826)
QY 1 MetaIaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSer 20
DB 37 ATGGACACCTCCAAAGTGAAGAGACACCTCGATCCCTGACGCGCTTGCAGCCCTCATCC 96
QY 21 SerGlnAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40
DB 97 TCAGAGCGCAGAGCCCTGCTGCTCTCTCTCCCGGGACCTGGGCTCCGCCAGCCA 156
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaCysAla 60
DB 157 TTGTCTATCTCTTTCGGGAGCTACTTGGCTGAGGATTTATGTGTGGAGCTCCCAAGGCC 215
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGlnAspPheSerCys 80
DB 216 TGTGGATTCCTCCCTGTTATCATTCGCTTTCGCTGCGACATGAGGACCTTCTCTTGC 275

```

QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
 |||||||  
 Db 276 TGGTTTCCCAACGACCATTTCTGCATAGAGACGTGGACACTCAAGCTTGGTCTAC 335  
 |||||||  
 QY 101 ArgLeuArgPheIlePheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
 |||||||  
 Db 336 AGGTACGCTTTATTTCCCTGCAGCTGGTTGGCTGGAGACATGTACACCTT-GGGCTG 394  
 |||||||  
 QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
 |||||||  
 Db 395 CGCAAGA-ATTGACCAAGTGCATCTTGACTTACATGTTTAGAACATCTTTGGCTAG 453  
 |||||||  
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
 |||||||  
 Db 454 CACCGCAGTGACCTGTGAGTGGGCGCTCCCGTGGGCTTACATGAAAGAGAGAGGA 513  
 |||||||  
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
 |||||||  
 Db 514 GAGTTCCTGACCTGGCGCTGCTGACTGGCCAGATGGCTGTGACAGAGCCAGCGC 573  
 |||||||  
 QY 181 ProGlyLeuLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200  
 |||||||  
 Db 574 CCAGGAGAGCTGCTGACAGCGGTGACATTCAAGCTCTGTGCGCCAGCGCTGGCGAT 633  
 |||||||  
 QY 201 ValIleGlnGlnAsnPheValThrArgArg 211  
 |||||||  
 Db 634 GTGATCCAGGCGCAGAACTTGTGACACGACGAG 666  
 |||||||  
 RESULT 9  
 LOCUS BF150250 675 bp mRNA linear EST 29-DEC-2000  
 DEFINITION uy83609.y1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:3666184 5'  
 similar to TR:P97423 P97423 JANUS KINASE 3 ;, mRNA sequence.  
 ACCESSION BF150250  
 VERSION BF150250.1 GI:11031645  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 675)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Other ESTs: uy83e09.x1  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapds-remail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 image.lnl.gov/image/html/iresources.shtml  
 MGI:1426952  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 390:  
 Location/Qualifiers  
 1. 675  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone IMAGE:3666184  
 /clone\_lib="NCI\_CGAP\_Mam5"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="7 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: Salt;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigators

BASE COUNT 138 a 215 c 197 g 123 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.62e-69 Length: 675  
 Score: 82.00 Matches: 170  
 Percent Similarity: 97.70% Conservative: 0  
 Best local Similarity: 97.70% Mismatches: 3  
 Query Match: 7.46% Indels: 4  
 DB: 12 Gaps: 0  
 US-09-397-967-16 (1-1099) x BF150250 (1-675)  
 QY 717 GllProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGln 736  
 |||||||  
 Db 37 GGACCGCCCAATCATTCTGCTGGAGCCCGCAAAAGCTGAAATTCTATGAGGACCA 96  
 |||||||  
 QY 737 GlyIleLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAla 756  
 |||||||  
 Db 97 GGACAGTGGCCCGCTCTCAATATGACAGACTGGGGGACTTATACACAGTGCATGGCG 156  
 |||||||  
 QY 757 TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle 776  
 |||||||  
 Db 157 TATGATCCTGGCGCGCGCTCTCGACTATCTCAAGACTTCAACGGCTCATTT 216  
 |||||||  
 QY 777 ThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGlu 796  
 |||||||  
 Db 217 ACATCAGATTACAGACTCTCTCAAGACCCACACTGGCATCCGAGTCCCTCGAGATGAG 276  
 |||||||  
 QY 797 LeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGlnArg 816  
 |||||||  
 Db 277 CTGTGCG--GT-GGCGCCCACTATGCTGATGCCAGAGCCCGCATATTCGAGAGAGA 333  
 |||||||  
 QY 817 HisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGlnLeuCysArg 836  
 |||||||  
 Db 334 CACCTTAAGTACATCTCTTCTGGGCAAGGCAACTTGGCAGCGTGGAGCTGTGCCG 393  
 |||||||  
 QY 837 TyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSer 856  
 |||||||  
 Db 394 TATGACCCCTGGGGGAGCAATATACGGGACCTGTGGGAGTGAAGAACTACAGACAGC 453  
 |||||||  
 QY 857 Val-ProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSe 876  
 |||||||  
 Db 454 GG-GCCAGACCAAGAGAGAGACTTCCAGCGGAGATTCAGATTCCTTAAGCTCTGCACAG 512  
 |||||||  
 QY 876 RasPheIleValLysTyrArgGlyValSerTyrGlyPro 889  
 |||||||  
 Db 513 CGACTTCATCGTCAAGTACCGGGGAGTCAAGCTATGGGCCA 552  
 |||||||  
 RESULT 10  
 LOCUS BG870513 907 bp mRNA linear EST 29-MAY-2001  
 DEFINITION 602791432F1 NCI\_CGAP\_S62 Mus musculus cDNA clone IMAGE:4922742 5',  
 mRNA sequence.  
 ACCESSION BG870513  
 VERSION BG870513.1 GI:14221053  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 907)  
 AUTHORS NIH-MGC http://mgc.ncbi.nlm.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapds-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L1AM10842 row: p column: 07  
High quality sequence stop: 742.

Location/Qualifiers

1. 907

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:4922742"

/clone\_lib="NCI CGAP\_SG2"

/lab\_host="DH10B (TI phage-resistant)"

/note="Organ: salivary gland; Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 166 a 279 c 275 g 185 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 2,926-64 Length: 907  
Score: 77.00 Matches: 77  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 7.01% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BG870513 (1-907)

QY 916 LeuH1stHspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeu 935

DB 216 GTCGACACGACCGCTACTGCTTCGCTGCGTGCACATCTGCAGGCGCATGGATCCTG 275

QY 936 G1yAlaArgArgCysValH1stArgAspLeuAlaAlaArgAsnIleLeuValGluSerGlu 955

DB 276 GGTGGCGCGCGCTGCTACACGACGACGCTGCGCGCGCAACATCTTGGTGGAGAGCGAG 335

QY 956 AlaH1stValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyr 975

DB 336 GCTCATGTGAAGATCCGCGGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 395

QY 976 TyrValValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGlu 992

DB 396 TACGTGTGCGCGACGCTGCGCAAGCCCATCTTTGGTATGCCCGGAG 446

RESULT 11

AA023709

LOCUS

AA023709 453 bp mRNA linear EST 21-JAN-1997

DEFINITION

mh77e02.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA

clone IMAGE:456986 5' similar to gb:L32955 Mouse protein tyrosine

kinase (MOUSE);, mRNA sequence.

ACCESSION

AA023709

VERSION

AA023709.1

KEYWORDS

EST.

SOURCE

ORGANISM

house mouse.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 453)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Gelsel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HM Mouse EST Project

Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the

IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

MGI:273874

Seq primer: -28M13 rev2 from Amer sham

High quality sequence stop: 406.

Location/Qualifiers

1. 453

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="IMAGE:456986"

/clone\_lib="Soares mouse placenta 4NbMP13.5 14.5"

/sex="unknown"

/tissue="placenta"

/dev\_stage="adult"

/lab\_host="DH10B"

/note="Organ: placenta; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer

15',

TGTTACCATCTGAAGTGGAGCGCGCGGCAATTTTCTTTTCTTTTCTTTT

T 3 1; double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 79 a 153 c 137 g 84 t

ORIGIN

Alignment Scores:  
Pred. No.: 1,116-61 Length: 453  
Score: 74.00 Matches: 118  
Percent Similarity: 98.33% Conservative: 0  
Best Local Similarity: 98.33% Mismatches: 1  
Query Match: 6.73% Indels: 2  
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AA023709 (1-453)

QY 957 H1stValLysIleAlaAspPheGlyLeuAlaLysLeuProLeuGlyLysAspTyr 976

DB 3 CATGTAAAGATCGCGAGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 62

QY 977 ValValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuAsp 996

DB 63 GTGTCGCGGAGCGCTGCGCAAGCCCATCTTTGGTATGCGCGGAGTCCATCTGAC 122

QY 997 AsnIlePheSerArgGlnSerAspValTTPSerPheGlyValValLeuTyrGluLeuPhe 1016

DB 123 AACATCTTCTCCCGCAACTGTGACGTGTGAGCTTCGAGTGTGTGTGACGACTCTTC 182

QY 1017 ThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArg-MetMetGlyProG 1036

DB 183 ACCTACTGCGACAGAGCTGACAGCCCATCCGCTGAGTCTTGAG-CATGATGGGGCTGA 241

QY 1036 uATGGLUGLProProLeuGlyAspArgLeuGluLeuAlaGluGlyArgLeuP 1056

DB 242 GCGTGAAGAGACCCCGGCTGCGCGCTGCGAGCTGCTGCGAGAGGCGGAGCGCTCC 301

QY 1056 oProProProThrcysProThrcyluValGlnGluLeuMetGlnLeuGlyTrrpAlaPro 1075

DB 302 ACCACCTCCACCTGCGCCGACGAGGTTCAGAGCTCATGACCTGTGCTGGCGGCC 359

RESULT 12

AA023670

LOCUS

AA023670 478 bp mRNA linear EST 21-JAN-1997

DEFINITION

mh78g06.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA

clone IMAGE:457114 5' similar to gb:L32955 Mouse protein tyrosine

kinase (MOUSE);, mRNA sequence.

ACCESSION

AA023670

VERSION

AA023670.1 GI:1487587

KEYWORDS

EST.

ORGANISM	Mus musculus
SOURCE	house mouse.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 478)
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE	The WashU-HMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Maria M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:274002 Seq primer: -28M13 rev2 from Amersham High quality sequence stop: 463. Location/Qualifiers
FEATURES	1..478 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:457114" /clone.lib="Soares mouse placenta 4NDMP13.5 14.5" /sex="unknown" /tissue_type="placenta" /dev_stage="adult" /lab_host="DH10B" /note="Organ: placenta; Vector: pT730-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCCGGAGAAATTTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT	83 a 160 c 142 g 93 t
ORIGIN	
Alignment Scores:	
Pred. No.:	1.18e-61 Length: 478
Score:	74.00 Matches: 118
Percent Similarity:	98.33% Conservative: 0
Best local Similarity:	98.33% Mismatches: 1
Query Match:	6.73% Indels: 2
DB:	9 Gaps: 0
US-05-397-967-16 (1-1099) x AA023670 (1-478)	
Oy	HISVALVLSITIEALASPPhEgLYeunAlALysLeuLeuProLeuGlyLysAspTyrTyr 976
Db	3 CATGGAAGATCGCGGACTTCGCTCGCTAGAGCTGTGCCCCCTGGGAAGACACTAC 62
Oy	977 VALVALAArgLuproGlyLInserProIlePheTyrPYrAlaProGlnSerLeuSerasp 996
Db	63 GTGGTCCCGAGACCTGGCCAAAGCCCAATCTTTGTGTATGCCCGGAGTCCCTACTGAC 122
Oy	997 AsnIlePheSerArgGlnSerAspYalTrrpSerPheGlyValValLeuTyrGluLeuPhe 1016
Db	123 AACACTCTTCCCGGCATCTGACCTGTGGAGCTTCGGAGAGTGCTGTAGAGACTTC 182
Oy	1017 ThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArg-MetMetGlyProG 1036
Db	183 ACCTACTCGCAGACAGCTGAGCCCATCCGCTGAGTTCCTGAG-CATGATGGGGCGCTGA 241

Qy	1036	uagcgluglucgProProLencCysArgLeuGluLeuLeuAaIagLgIYAArgLeuP	1056
Db	242	GGCTGAAGAGACCCCGCTCTGCCGCTCTCTGGACGCTGCTGGACAGAGGCGCACGCTCC	301
Qy	1056	oProProProthrhCysProThrgluValglnGluLeuMetglnLeuCysTrpAlaPro	1075
Db	302	ACCACCTCCACCTGCGCCACCGAGGTTCAGAGCTCATGACGTGTGTGTGGCGGCC	359
RESULT 13			
LOCUS	AA881653	593 bp	mRNA
DEFINITION	vx21c02.r1 Soares_thymus_2NDMT Mus musculus cDNA clone		EST 26-MAR-1998
ACCESSION	AA881653		
VERSION	AA881653.1		
KEYWORDS	GI:2990963		
ORGANISM	house mouse.		
SOURCE	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 593) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheinberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.		
JOURNAL	The Washn-HM Mouse EST Project		
COMMENT	Unpublished (1996) Contact: Marra M/Mouse EST Project Washn-HM Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:667642 Seq primer: -28m13 rev2 Err from Amersham High quality sequence stop: 556. Location/Qualifiers 1..593 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone IMAGE:1265090" /clone_idb="Soares_thymus_2NDMT" /sex="male" /tissue_type="Thymus" /dev_stage="4 weeks" /lab_host="DH10B" /note="vector: pUT73P-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5' TGTACCAATCATGAGAGCGAGCGCGCGCTGTTTATTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Hento Soares and M.Patima Bonaldo."		
BASE COUNT	106 a	166 g	124 t
ORIGIN	197 c		
Alignment Scores:	1.45e-60	Length:	593
Pred. No.:	73.00	Matches:	117
Score:	98.32%	Conservative:	0
Percent Similarity:	98.32%	Mismatches:	1
Best Local Similarity:	6.64%	Indels:	2
Query Match:	9	Gaps:	0
DB:			

US-09-397-967-16 (1-1099) x AA881653 (1-593)

QY 958 VALVLLIIEALAApPhgLyLeuAlaLysLeuProLeuGlyLysAspTyrVal 977  
 DB 15 GTGAAGATCGCGGACTTCGGCTCTAGCTGTGCCCCCTGGAGAAAGCACTACTACGTG 74

QY 978 VALATGluProGluGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAsn 997  
 DB 75 GTCCCGAGACCTGGCCCAAGGCCCATCTTTGGTATGGCCCGAGAGCTTATCTGCAAC 134

QY 998 IIEPheSerArgGlnSerAspValTyrPheSerPheGlyValValLeuTyrGluLeuPheThr 1017  
 DB 135 ATCTTCTCCCGCCCAATCTGACGTGTGAGCTTCGGAGTGGTGTGTGACAGCTCTTCACC 194

QY 1018 TyrCysAspLysSerCysSerProSerAlaGluPheLeuArg-MetMetGlyProGluArg 1037  
 DB 195 TACTCGACAAAGAGCTGCGAGCCCATCTGCTGAGTTCCTGAG-CATGATGGGCGCTGAGCG 253

QY 1037 GGLuGlyProProLeuGlyArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPr 1057  
 DB 254 TGAAGAGACCCCGCTCTGCGGCTCTCTGAGCTGTGAGAGGAGGCGGAGCGCTCCACC 313

QY 1057 oProProThrCysProThrGluValGlnLeuLeuMetGlnLeuCysTyrAlaPro 1075  
 DB 314 ACCTCCACCTGCGCCACCGAGGTTGAGAGCTCATGACAGCTGTGCTGGCGGCC 368

RESULT 14  
 LOCUS BM935182 698 bp mRNA linear EST 13-MAR-2002  
 DEFINITION UI-M-BH3-agz-f-08-0-UI-r1 NIH\_BMAP\_M\_S4 Mus musculus cDNA clone  
 ACCESSION BM935182  
 VERSION BM935182.1 GI:19394334  
 KEYWORDS EST  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 698)  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: Chin, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mes@nhi.nih.gov  
 CDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 REVERSE.  
 FEATURES  
 source location/Qualifiers  
 1..698  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-BH3-agz-f-08-0-UI"  
 /clone\_lib="NIH\_BMAP\_M\_S4"  
 /dev\_stage="27-32 days"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; The  
 NIH\_BMAP\_M\_S4 library is a subtracted library of a series,  
 ultimately derived from a mixture of individually tagged  
 normalized libraries from ten regions of the mouse brain  
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,

cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH\_BMAP\_M\_S4, NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, NIH\_BMAP\_M\_S3.1, NIH\_BMAP\_M\_S2, NIH\_BMAP\_M\_S1. The subtracted library (NIH\_BMAP\_M\_S4) was constructed as follows: PCR amplified cDNA inserts from NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH\_BMAP\_M\_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)"

BASE COUNT 132 a 211 c 215 g 138 t 2 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,66-58 Length: 698  
 Score: 71.00 Matches: 183  
 Percent Similarity: 95.31% Conservative: 0  
 Best Local Similarity: 95.31% Mismatches: 3  
 Query Match: 6.46% Indels: 9  
 DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x BM935182 (1-698)

QY 801 GLVALAGlnLeuTyrAlaCysGlnAspProAlaIlePheGlnGluArgHisLeuTyr 820  
 DB 27 GGGGCCACCTGATGCTCTGCGAGACCCCGCATATTCAGAGAGACACCTTAAGTAC 86

QY 821 IIESerLeuGlnGlyLysGlnAsnPheGlySerValGlnLeuGlnCysArgTyrAspProLeu 840  
 DB 87 ATCTCTTCTGCGCAAGGCAACTTTGGACGCTGACCTGTGCGCTATGACCCCTG 146

QY 841 GLYAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerVal-ProAspG 860  
 DB 147 GGGGACAAATACGGGAGACCCCTGCGAGCAATGAAACACTACAGACACGGG-GCCAAACCA 205

QY 860 nGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheLeu 880  
 DB 206 GCAGAGGACCTTCCAGCGGAGATTCAGATCTTAAGGCTCTGCACAGGACTTCATCGT 265

QY 880 IlySTyrArgGlyValSerTyrGlyProGluArgGlnSerLeuArgLeuValMetGluTyr 900  
 DB 266 CAAGTACCGGGAGTACGATATGGCCAGGCTGCCAGAGCTCGGTGGTATGGAGTGA 325

QY 900 IIEuProSerGlyCysLeuArgAspLeu-LeuGlnArgHis----ArgGlyLeuHisThr 918  
 DB 326 CTGCGCCAGCGGCTGCTGGGAGA-CTTCTGAGAGGCCATGCCGGCGCG-CGTCAACACC 383

QY 919 AspArgLeuLeuLeuPheAlaTyrGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg 938  
 DB 384 GACCGCTACTGCTGTTCGCTTGGAGATCTCAAGGAGCATGAGATACCTGGCGCGCC 443

QY 939 ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHisVal 958  
 DB 444 CGCTGGGTACACCGTACCGTGGCGGCAACATCTTGATGAGAGCGGAGCTCAATGTC 503

QY 959 LysIIEAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal 978  
 DB 504 AAGATCGCGGACTTCGGCTCGCTAAGCTCTGCCCTGGGAAAGCACTACGTGTGTC 563

QY 979 ArgGluProGlyGlnSerProIle 986  
 DB 564 CGGAGCTGCGCAAGGCCCATC 587

RESULT 15  
BF181053  
LOCUS  
DEFINITION BF181053 344 bp mRNA linear EST 31-OCT-2000  
601805105f1 NCL.CGAP\_Mam5 Mus musculus cDNA clone IMAGE:4035936 5',  
mRNA sequence.  
ACCESSION BF181053  
VERSION BF181053.1 GI:11059195  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 344)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapsb@mail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LLM9310 row: j column: 01  
High quality sequence stop: 344.  
Location/Qualifiers  
1..344  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_image="IMAGE:4035936"  
/clone\_lib="NCI CGAP Mam5"  
/tissue\_type="tumor\_gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"

BASE COUNT 73 a 95 c 108 g 68 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6,27e-57 Length: 344  
Score: 69.00 Matches: 113  
Percent Similarity: 97.41% Conservative: 0  
Best Local Similarity: 97.41% Mismatches: 2  
Query Match: 6.28% Indels: 3  
DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BF181053 (1-344)  
QY 646 ArglyValLeuLeuAlaArgGlyGlyAspGlyAsnPropropheileLeuSer 665  
DB 1 CGGAAGCTGCTCTCGCTCGTGAAGGGGAGTGAATGCACCTTCATTAACTGAGT 60  
QY 666 AspProGlyValSerProThrValLeuSerLeuLettLeuThrAspArgIleProTyr 685  
DB 61 GATCTGTGTGTCATCCACAGCTGTCTGAGCTGGAATGCTCACCGACAGATVACCTGC 120  
QY 686 ValAlaProGlyLeuGlnGluAlaGlnThrLeuGlyLeuGlnAlaAspGlyTyrGly 705  
DB 121 GTGGCCCCCAATGCTCCAGAGAGCTCAGACACTGCTTTGGAGCTGACAAAGTGGGCC 180  
QY 706 PheGlyAlaThrThrTyrGluValPheGlnArg-GlyProAlaHisIleThrSerLeuGln 725  
DB 181 TTGGAGCCCAACAGCTGGAGGAGTGT-CAGCGGGGAGCCGCCACATCACTCGCTGGA 239  
QY 725 uProAlaLysLysLeuysPheTyrGluAspGlnGlyLeuProAlaLeuysTyrPth 745

DB 240 GCCGCCAAGAAAGCTGAAGTTCTATGAGGACA-GGACAGCTGCCCTCTCAATGAGC 298  
QY 745 rgLeuLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760  
DB 299 AGAAGTGGCGGACTTATCACACAGTGCATGCGCTATGATCTCTGCC 344  
RESULT 16  
AM141115  
LOCUS  
DEFINITION AM141115 436 bp mRNA linear EST 30-OCT-1999  
RG1BC74 5' end similar to protein-tyrosine kinase jak3, mRNA  
sequence.  
ACCESSION AM141115  
VERSION AM141115.1 GI:6160891  
KEYWORDS EST.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 436)  
AUTHORS Lee,N.H., Glodex,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
Kerlavage,A.R. and Adams,M.D.  
TITLE Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat  
Gene Index  
JOURNAL Unpublished (1998)  
COMMENT Contact: Lee, NH  
The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@igrr.org  
For clone availability, additional sequence and expression  
information related to this EST please check the TIGR Rat Gene  
Index (<http://www.tigr.org/tdb/ryl/ryl.html>). To order a clone  
contact the ATCC (<http://www.atcc.org/atcc.html>).  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1..436  
/organism="Rattus sp."  
/db\_xref="taxon:10118"  
/clone="RG1BC74"  
/clone\_lib="Normalized rat embryo, Bento Soares"  
/dev\_stage="embryo 8, 12, 18 dpc"  
/note="Vector: pT7mpac; Site\_1: EcoRI; Site\_2: NotI"

BASE COUNT 72 a 134 c 108 g 122 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.84e-53 Length: 436  
Score: 65.00 Matches: 65  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.91% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AM141115 (1-436)  
QY 13 GlnArgSerCysSerLeuSerSerGlnAlaGlyAlaLeuHisValLeuLeuProPro 32  
DB 61 CAGGCTCTTGTATACCTCTATCTCAGAAAGCAGACCCCTGCATGCTGCTTCCCTCC 120  
QY 33 ArgGlyProGlyPropProGlnArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAsp 52  
DB 121 CGGGAGACTGGGCTCCCGACGACTGTCTCTTTTGGGAGTACTTGGCTGAGGAT 180  
QY 53 LeuGlyValArgAlaAlaLysAlaCysGlyIleLeuProValTyrHisSerLeuPheAla 72  
DB 181 CTGTGTGTGGGGGTGCCAAGGCTGTGGCATCTGCGCCCTTATATGATTCGCTTCGT 240  
QY 73 LeuAlaThrGluAsp 77  
DB 241 CTGGCCACCGAGGAC 255





```

/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone_image="IMAGE:5003274"
/clone_lib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"

```

```

/organism="Mus musculus"
/strain="C3H x 101 F1"
/db_xref="taxon:10090"
/clone="IMAGE:2780302"

```

ACCESSION AA755769

```

VERSION      AA755769.1  GI:2802967
LOCUS        EST.
KEYWORDS     house mouse.
SOURCE       Mus musculus
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 867)
AUTHORS      Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
              Geisels, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
              Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,
              Treising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
              Waterston, R.
TITLE        The WashU-HMI Mouse EST Project
JOURNAL      Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
              WashU-HMI Mouse EST Project
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: mouseest@watson.wustl.edu
              This clone is available royalty-free through LLNL; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              MGI:649972
              Seq primer: -26m13 rev1 ET from Amersham
              High quality sequence stop: 412.
              Location/Qualifiers
                1..867
                 /organism="Mus musculus"
                 /strain="NIH Swiss"
                 /db_xref="taxon:10090"
                 /clone_image="1224380"
                 /clone_lib="Stratagene mouse heart (#937316)"
                 /sex="pooled"
                 /tissue_type="heart"
                 /dev_stage="13 day embryos"
                 /lab_host="SOLR (kanamycin resistant)"
                 /note="Organ: heart; Vector: pBluescript SK-; Site.1:
                 EcORI; Site.2: XhoI; Cloned unidirectionally. Primer:
                 Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts.
                 Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
                 adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor
                 sequence: 5' CTCGAGCTTTT TTTT TTTT TTTT 3'"
BASE COUNT   179 a      239 c      292 g      157 t
ORIGIN
Alignment Scores:
Pred. No.:   7,87e-39
Score:       51.00      Matches:      867
Percent Similarity: 97.50%      Conservative: 0
Best Local Similarity: 97.50%      Mismatches: 2
Query Match: 4.64%      Indels:      2
DB:          9          Gaps:      0
US-09-397-967-16 (1-1099) x AA755769 (1-867)
QY 991 ProGUlSerLeuSerAspAsnIlePheSerArgInSerAspValIrrPserPheGlyVal 1010
    |||||||
Db 454 CCGAGATGCCATATCTGACACATCTCTCCGCCCATCTGACGTGGAGCTTGGAAGTG 395
QY 1011 ValLeuTyrGluLeuPheTrpTyrCysAspIlySerCysSerProSerAlaGluPheLeu 1030
    |||||||
Db 394 GTGTTGTACAGAGCTTCTACCTACTGCGACAGAGCTGCAGCCCATCGCTGAGTTCCTG 335
QY 1031 ArgMetMetGlyProGUATrGAGGluGlyProProLeuCysArgLeuLeuGluLeuAla 1050
    |||||||
Db 334 CGCATGTATGGGGCTTGACCTGAGAGACCCCGTC-TGCCGCTCTCTGAGACTCTGGCA 276
QY 1051 GluGlyArgArgLeuProProProProThrCysProThrGluValGluGluMetGln 1070
    |||||||
Db 275 GA-GGGCGAGCGCTCCACACACCTCCACCTGCCCGCCAGCGATTCAAGAGCTATCGAG 217
RESULT 23

```

```

BE753801
LOCUS        BE753801      543 bp      mRNA      linear      EST 25-APR-2001
DEFINITION   206963 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION    BE753801
VERSION      BE753801.1  GI:10167793
KEYWORDS     EST.
SOURCE       cow.
ORGANISM     Bos taurus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
              Bovidae; Bovinae; Bos.
REFERENCE    1 (bases 1 to 543)
AUTHORS      Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
              Casas, E., Wray, J.E., White, J., Cho, J., Fahrnenkrug, S.C., Bennett,
              G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
              Perle, G., Holt, I., Karanmycheva, S., Liang, F., Quackenbush, J. and
              Kelle, J.W.
TITLE        Sequence evaluation of four pooled-tissue normalized bovine cDNA
              libraries and construction of a gene index for cattle
JOURNAL      Genome Res. 11 (4), 626-630 (2001)
COMMENT      Contact: Smith TPL
              USDA, ARS, US Meat Animal Research Center
              PO Box 166, Clay Center, NE 68933-0166, USA
              Tel: 402 762 4366
              Fax: 402 762 4390
              Email: smith@email.marc.usda.gov
              Single pass sequencing. Bases called and alt trimmed with phred
              v0.980904.e. Vector identified by cross_match with the -mnscore 18
              and -primatch 12 options.
              PCR Primers:
              FORWARD: AGGAACAGCTATGACCAT
              BACKWARD: GTTTCACGTCACGACG
              Plate: 49 row: M column: 16
              Seq primer: ATTAGGTGACACTATAG.
              Location/Qualifiers
                1..543
                 /organism="Bos taurus"
                 /db_xref="taxon:9913"
                 /clone_lib="MARC 2BOV"
                 /tissue_type="pooled"
                 /lab_host="DH10B"
                 /note="Vector: pCMV SPORT6; Site.1: NotI; Site.2: SalI;
                 Library made from pooled tissue from testis, thymus,
                 semitendinosus muscle, longissimus muscle, pancreas,
                 adrenal, and endometrium."
BASE COUNT   100 a      182 c      158 g      103 t
ORIGIN
Alignment Scores:
Pred. No.:   4.1e-37
Score:       49.00      Matches:      543
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 4.46%      Indels:      0
DB:          12          Gaps:      0
US-09-397-967-16 (1-1099) x BE753801 (1-543)
QY 860 GInGlnArgAspPheGlnArgGluIleGlnIleLeuAlaLeuHisSerAspPheIle 879
    |||||||
Db 12 CAGCAGAGGGACCTTCAGCGGAGATCCAGATCTCTCAAGCCCTCCACAGTGAATTCATT 71
QY 880 ValIlyTyrArgGlyValSerTyrGlyProGUATrGAGGluGlyProProLeuCysArgLeuValMetGln 899
    |||||||
Db 72 GTCAAGTACCGGGGTGTACGCTATGCCCCAGGCCGCCAGAGCTTGGTATGAGAG 131
QY 900 TyrLeuProSerGlyCysLeuArgAsp 908
    |||||||
Db 132 TATCTGCCCAAGTGGCTGCTGCGCGAC 158
RESULT 24
BI339676

```

LOCUS BI339676 557 bp mRNA linear EST 30-JUL-2001  
 DEFINITION 364724 MARC 2Pig Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION BI339676  
 VERSION BI339676.1 GI:15032955  
 KEYWORDS EST.  
 SOURCE Pig.  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 REFERENCE 1 (bases 1 to 557)  
 AUTHORS Fahrnkruug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,  
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.M.  
 and Keeler,J.W.  
 TITLE Design and use of two pooled tissue normalized cDNA libraries for  
 EST discovery in swine  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smtlh@mail.marc.usda.gov  
 Single pass sequencing. Bases called and alt\_trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -mnscore 18  
 and -mismatch 12 options.  
 PCR primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCCTCAGTCACGACG  
 Plate: 102 row: E column: 17  
 Seq primer: ATTAGGTGACACTATAG.  
 FEATURES  
 source  
 1..557  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /clone\_lib="MARC 2Pig"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
 library made from pooled tissue from testis, ovary,  
 endometrium, hypothalamus, pituitary, and placenta."  
 BASE COUNT 115 a 185 c 159 g 98 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.22e-37 Length: 557  
 Score: 49.00 Matches: 49  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.46% Indels: 0  
 DB: 13 Gaps: 0  
 US-09-397-967-16 (1-1099) x BI339676 (1-557)  
 QY 860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879  
 Db 238 CACACAGAGGACTTCAGCGGGAGATCCATTAAGCCCTCCACAGCAGCACTTCATT 297  
 QY 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899  
 Db 298 GTCACAGTACCGGGGTCTACGATATGCCCCAGCGCCCGCAGAGTCTCGGGCTGTCATGGAA 357  
 QY 900 TyrLeuProSerGlyCysLeuArgAsp 908  
 Db 358 TATCTGCCAAGCGGCTGCTCGGGGAC 384  
 RESULT 25  
 LOCUS BF705690 544 bp mRNA linear EST 29-DEC-2000  
 DEFINITION 243606 MARC 2Pig Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION BF705690  
 VERSION BF705690.1 GI:11997351  
 KEYWORDS EST.  
 SOURCE Pig.

ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 REFERENCE 1 (bases 1 to 544)  
 AUTHORS Fahrnkruug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,  
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.M.  
 and Keeler,J.W.  
 TITLE Design and use of two pooled tissue normalized cDNA libraries for  
 EST discovery in swine  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smtlh@mail.marc.usda.gov  
 Single pass sequencing. Bases called and alt\_trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -mnscore 18  
 and -mismatch 12 options.  
 PCR primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCCTCAGTCACGACG  
 Plate: 70 row: D column: 5  
 Seq primer: ATTAGGTGACACTATAG.  
 FEATURES  
 source  
 1..544  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /clone\_lib="MARC 2Pig"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
 library made from pooled tissue from testis, ovary,  
 endometrium, hypothalamus, pituitary, and placenta."  
 BASE COUNT 96 a 181 c 156 g 111 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.55e-34 Length: 544  
 Score: 46.00 Matches: 46  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.19% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-397-967-16 (1-1099) x BF705690 (1-544)  
 QY 973 LysAspTyrTyrValValArgGlnProGlyGlnSerProIlePheTyrTyrAlaProGlu 992  
 Db 175 AAAGACTACTATGTGTCTCGGACGCCGCCAGAGCCCATCTCTGCTATGCCCGGAG 234  
 QY 993 SerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrPheGlyValValLeu 1012  
 Db 235 TCCCTCTAGACATATCTTCTCCGCCAGTCGATCTTGAGCTTGGGGGTCTGCTG 294  
 QY 1013 TyrGlnLeuPheTyrTyr 1018  
 Db 295 TATGAGCTCTTACACCTAC 312  
 RESULT 26  
 LOCUS BF193178 562 bp mRNA linear EST 02-NOV-2000  
 DEFINITION 244670 MARC 2Pig Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION BF193178  
 VERSION BF193178.1 GI:11076547  
 KEYWORDS EST.  
 SOURCE Pig.  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 REFERENCE 1 (bases 1 to 562)  
 AUTHORS Fahrnkruug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,  
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.M.

TITLE design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine

JOURNAL Unpublished (2000)

COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.

PCR primers  
FORWARD: AGGAACAGCTATGACAT  
BACKWARD: GTTTCCAGTCACGACG  
Plate: 72 row: F column: 7  
Seq primer: ATTTAGTGACACTATAG.  
Location/Qualifiers  
1. 562  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_lib="MARC\_2P1G"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI; library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 97 a 188 c 167 g 110 t

ORIGIN

Alignment Scores: 3.69e-34 Length: 562  
Score: 46.00 Matches: 46  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.19% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BF193178 (1-562)

QY 973 LYSASPTTYTYRVAIYALARGIUPROGLYGINSERPROILEPHERTRYALAPROGLU 992  
|||||

DB 186 AAGAGTACTATGTGTGTCGCGAGCGCGCCAGCCCATCTTGTGATATCCCCGGAG 245  
|||||

QY 993 SerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeu 1012  
|||||

DB 246 TCCCTCTCAGACAATATCTTCTGCGCGCAGTCGATGTCTGGAGCTTCGGGCTCTGCTG 305  
|||||

QY 1013 TYRGIUleupherThyr 1018  
|||||

DB 306 TATGAGCTCTTCACCTAC 323

RESULT 27  
A2956543 659 bp DNA linear GSS 27-APR-2001  
LOCUS A2956543/c  
DEFINITION clone UUGC2M0223004 F, DNA sequence.  
ACCESSION A2956543  
VERSION A2956543.1 GI:13827770  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 659)  
Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weis,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0223 row: O column: 04  
Seq primer: GGTGTAAACGACGCGCAGT  
Class: plasmid ends  
High quality sequence stop: 659.

FEATURES  
source  
1. 659  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_lib="UUGC2M0223004"  
/clone\_lib="Mouse 10kb plasmid UUGC2M library"  
/sex="Female"  
/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42ov; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The ligated DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g147321419b1/AF129072.1), a copy number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 155 a 217 c 178 g 109 t

ORIGIN

Alignment Scores: 4.25e-33 Length: 659  
Pred. NO.: 45.00 Matches: 45  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 4.09% Gaps: 0  
DB: 17

US-09-397-967-16 (1-1099) x A2956543 (1-659)

QY 325 GIUALGIupheProGLYleupROGLUALeuserPheValAlaleuValaspGLYTYR 344  
|||||

DB 452 GAAGCGGAGTTTCCGGGGCTGCTGAGCGCTGCTTCGCGCCCTTGAGATGGGTAC 393  
|||||

QY 345 PheArgLeuIleCYASpSerArgHisTyrPheCysIysGluValAlaProProArgLeu 364  
|||||

DB 392 TTCCCGCTGATCTCGACCTCAGGCAATTAATTTCTGCAAGAGGTGGCGCCCGCGGCTG 333  
|||||

QY 365 LeuGIuGIuGIuAla 369  
|||||

DB 332 CTGGAGAGGAGGCG 318

RESULT 28  
B1540725 417 bp mRNA linear EST 30-AUG-2001  
LOCUS B1540725  
DEFINITION 454105 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.  
ACCESSION B1540725  
VERSION B1540725.1 GI:15381837  
KEYWORDS EST.

SOURCE	COW.
ORGANISM	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
AUTHORS	1 (bases 1 to 417) Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-Mckown,C.G., Pierce,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.
TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL	Genome Res. 11 (4), 626-630 (2001)
MEDLINE	21180013
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross-match with the -m1nscore 18 and -mismatch 12 options. PCR Primers FORWARD: AGGAACAGCTATGACCAT BACKWARD: GTTTTCCAGTCACGACG Plate: 118 row: P column: 4 Seq primer: ATTGAGTGACACTATGAC
FEATURES	Location/Qualifiers 1..417 /organism="Bos taurus" /db_xref="taxon:9913" /clone_lib="MARC 1B0V" /tissue_type="pooled" /lab_host="DH10B" /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI; Library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."
BASE COUNT	74 a 138 c 121 g 83 t 1 others
ORIGIN	
Alignment Scores:	
Pred. No.:	2.35e-32
Score:	44.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	4.00%
DB:	13 Gaps: 0
US-09-397-967-16 (1-1099) x B1540725 (1-417)	
OY	tyrtyrvalalarglunproglnglnserproilaepttyrtyralaproglnserleu 994 
Db	112 TACTACGTGGTCCGCGACGCCAGGACGCCCATCTTCTGTGTCGCCGAGTCCCTC 171 
OY	seraspanilipheserfarglnseraspaltrpserphegylvalleuylrclu 1014 
Db	172 TCGGACACATCTTCTCGCGCAGTCGATGTGGACCTTCGGGGTGTCTGTACGAG 231 
OY	leuphethtyr 1018 
Db	232 CTCCTTCACCTAC 243
RESULT 29	
AA020387	AA020387 441 bp mRNA linear EST 21-JAN-1997
LOCUS	
DEFINITION	mb60c03.r1 Soares mouse placenta 4bM213.5 14.5 Mus musculus cDNA clone IMAGE:455332.5' similar to gb:LA0172 Mus musculus JAK3 gene, complete cds (MUSM5)., mRNA sequence.
ACCESSION	AA020387
VERSION	AA020387.1 GI:1484169
FEATURES	EST.

SOURCE	mouse muscle.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus. 1 (bases 1 to 441)
AUTHORS	Mairal,M., Hillier,L., Allen,M., Bowles,M., Dierich,N., Dubouque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,S., Soares,B., Wilson,R. and Waterston,R.
TITLE	The WashU-HMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Maria M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LTLN; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:272220
FEATURES	Putative full length read vector to vector length is 448 Seq primer: -28M13 rev2 from Amersham High quality sequence stop: 208. Location/Qualifiers 1..441 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone_image="IMAGE:455332" /clone_lib="Soares mouse placenta 4NBMPJ.5 14.5" /sex="unknown" /tissue_type="Placenta" /dev_stage="adult" /lab_host="DH10B" /note="Organ: Placenta; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15'.
BASE COUNT	81 a 142 c 124 g 94 t
ORIGIN	GTTACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTTTT T 3'1: double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
Alignment Scores:	
Pred. No.:	2,51e-32
Score:	44.00
Percent Similarity:	100.00%
Best local Similarity:	100.00%
Query Match:	4.00%
DB:	9
US-09-397-967-16 (1-1099) x AA020387 (1-441)	
QY 1032 MetMetGlyProGluArgGluGlyProProLeuGlySargLeuGluLeuAlaGlu 1051	
Db 80 ATGATGGGGCGCTGAGCGTGAAGAGACCCCGCTCTGCGCCCTCTGGAAGCTGTCGAGAG 139	
QY 1052 GlyArgGluLeuProProProProThCysProThArgGluValGlnGluMetGlnLeu 1071	
Db 140 GGGCGAGCCCTCCACACACCTCCACACCTGCGCCACAGGTTACGGAGCTCATGACACTG 199	
QY 1072 CysTrpAlaPro 1075	
Db 200 TGTGTGGCGGCC 211	
RESULT 30	





DB:	12	Gaps:	0
US-09-397-967-16 (1-1099) x BF149581 (1-581)			
Oy	989	TyralaproguslerseusAspn1lepheserarginsersapvaltrpserpne	1008
Db	363	TATGCCCCGAGTCCCTATCTGACACATCTTCTCCGCCCAATCTGACGTGAGACTTC	422
Oy	1009	GLyAlValleuYrGluLeuPheRthYrCYaSpLySerCYsseRProSeRAlaGlu	1028
Db	423	GGAGTGTGTGTGATGAGAGCTCTTCACCTACTGCGACAGAGCTGACGCCATCCGCTGAG	482
Oy	1029	PheluN 1030	
Db	483	TTTCCTG 488	
RESULT 33			
LOCUS	BF191625	554 bp	linear
DEFINITION	239261 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.		EST 02-NOV-2000
ACCESSION	BF191625		
VERSION	BF191625.1	GI:11074994	
KEYWORDS	EST.		
SOURCE	pig.		
ORGANISM	Sus scrofa		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
AUTHORS	1 (bases 1 to 554) Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.M. and Keele,J.W.		
TITLE	Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4365 Fax: 402 762 4390 Email: smithemall.marc.usda.gov Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options. PCR Primers FORWARD: AGGAACAGCTATGACCAT BACKWARD: GTTTCCTCATCTGACGACG Plate: 67 row: B column: 8 Seq primer: ATTGAGTGACACTATGAG.		
FEATURES	Location/Qualifiers source 1..554 /organism="Sus scrofa" /db_xref="taxon.9823" /clone_lib="MARC 2P1G" /tissue_type="pooled" /lab_host="DH10B" /note="Vector: PCMV SPOT6; Site_1: NotI; Site_2: SalI; library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."		
BASE COUNT	121 a 176 c 150 g 107 t		
ORIGIN			
Alignment Scores:			
Pred. NO.:	2.85e-29	Length:	554
Score:	41.00	Matches:	41
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.73%	Indels:	0
DB:	12	Gaps:	0
US-09-397-967-16 (1-1099) x BF191625 (1-554)			
Y	860	GInGlnArGspPheGlnArGGuInleGInleuLYsAlaLeuHissersAspPhele	870

Db 432 CAGCAGAGGACCTTCACGCGGAGATCCAGATCCTTAAGCCCTCCACAGCAGCTTCAAT 491

Qy 880 VallystYrArglyValSerTyrglyPrgOglYArgInSerleuArgleuValMetGlu 899  
|||||

Db 492 GTCAAGTACCGGGGTGTACAGTATGGCCAGCGCCAGCTGTGGGGCTGTCAATGGA 551

Qy 900 Tyr 900  
|||

Db 552 TAT 554

RESULT 34  
AA920299 731 bp mRNA linear EST 20-APR-1998  
LOCUS AA920299  
DEFINITION v93002.r1 Soares.thymus.2NDMT Mus musculus cDNA clone  
IMAGE:1282802.5, similar to gb:L40172 Mus musculus JAK3 gene,  
complete cds (MOUSE);, mRNA sequence.

ACCESSION AA920299  
VERSION AA920299.1 GI:3067078  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 731)  
AUTHORS Mastra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
The WashU-HM1 Mouse EST Project  
Unpublished (1996)  
Contact: Mastra M/Mouse EST Project  
WashU-HM1 Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:674602  
Seq primer: -28m13 rev2 ET from Amersham  
High quality sequence stop: 215.  
Location/Qualifiers  
1. 731  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1282802"  
/clone\_lib="Soares\_thymus\_2NDMT"  
/sex="male"  
/tissue\_type="thymus"  
/dev\_stage="4 weeks"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAAGTGGAGCGCGGGAATTTTTTTTTTTTTTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. RNA  
provided by Dr. Bertrand Jordan. Library went through two  
rounds of normalization, and was constructed by Bento  
Soares and M.Fatima Bonaldo."

BASE COUNT 167 a 209 c 214 g 141 t

ORIGIN

Alignment Scores:  
Pred. No.: 3.59e-27 Length: 731  
Score: 39.00 Matches: 39  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.55% Indels: 0

DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AA920299 (1-731)

Qy 720 HisIleHnSerleuGluProAlaLysLysleuLysPheTyrgluAspGlnGlnleu 739  
|||||

Db 187 CACATCACTGCTGAGAGCCGCCAAAGCTGAAGTCTTGTAGGAGCAGGAGCAGCTG 246

Qy 740 ProAlaLeuLysTrpThgGluLeuIleThgGlnCysMetAlaTyraSp 758  
|||||

Db 247 CCGGCTTCATAATGACAGACTGGCGGACTTATACACACATGATGCTATGAT 303

RESULT 35  
AA023069 410 bp mRNA linear EST 21-JAN-1997  
LOCUS AA023069  
DEFINITION mh6d02.r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA  
clone IMAGE:455907.5, similar to gb:L32955 Mouse protein tyrosine  
kinase (MOUSE);, mRNA sequence.

ACCESSION AA023069  
VERSION AA023069.1 GI:1487776  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 410)  
AUTHORS Mastra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
The WashU-HM1 Mouse EST Project  
Unpublished (1996)  
Contact: Mastra M/Mouse EST Project  
WashU-HM1 Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:272795  
Seq primer: -28M13 rev2 from Amersham  
High quality sequence stop: 299.  
Location/Qualifiers  
1. 410  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:455907"  
/clone\_lib="Soares mouse placenta 4NDMP13.5 14.5"  
/sex="unknown"  
/tissue\_type="placenta"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Organ: placenta; Vector: pT73D-Pac (Pharmacia)  
with a modified polylinker; Site.1: Not I; Site.2: Eco RI;  
1st strand cDNA was primed with a Not I - oligo(dT) primer  
[5'  
TGTTACCAATCTGAAGTGGAGCGCGGGAATTTTTTTTTTTTTTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 67 a 152 c 128 g 62 t 1 others

ORIGIN

Alignment Scores:  
Pred. No.: 1.72e-26 Length: 410  
Score: 38.00 Matches: 38  
Percent Similarity: 100.00% Conservative: 0

US-09-397-967-16 (1-1099) x AAL174246 (1-308)

OY 916 leuHSTHRspARgLeuLeuLeuLeuLeuLeuLeuLaATrrpgInllEcysLysGlyMetcIduryleu 935  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 20 CTGCACACCAGCAGCGCTTCTGTGGCTTGCCATTGCCAGACTTCGCAAGGCCATGGAGTACCTG 79

OY 936 GlAAlArGrCySValHisrAsrpLeuAlaAlArGsntle 950  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 80 GGTGGCGCCGGCTGCTACACCGTTACCTGCTGGTGGCGCACATC 124

RESULT 37  
AM501263 231 bp mRNA linear EST 01-MAR-2000  
LOCUS AM501263  
DEFINITION UI-HF-BPDP-au-d-02-0-UI.r1 NIH\_MGC\_51 Homo sapiens cDNA clone  
IMAGE:3073082 5', mRNA sequence.  
ACCESSION AM501263  
VERSION  
KEYWORDS EST.  
SOURCE AWS01263.1 GI:7114630  
ORGANISM human.  
Human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: rcgabs@remail.nih.gov  
ECO RI site shown at the beginning of the sequence.  
Tissue procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: M.B. Soares Lab  
CDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing Arranged by: M.B. Soares Lab  
clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
[www.bio.lnl.gov/bdrip/image/image.html](http://www.bio.lnl.gov/bdrip/image/image.html)  
Seq primer: M13 Forward.

FEATURES  
source Location/Qualifiers  
1..231  
    /organism="Homo sapiens"  
    /db\_xref="taxon:9606"  
    /codon\_table="IMAGE:3073082"  
    /clone\_id="NIH\_MGC\_51"  
    /tissue\_type="lymph"  
    /cell\_line="germinal center B cells"  
    /lab\_host="DH10B (LT1)"  
    /note="Vector: pRT73-Pac; Site\_1: NotI; Site\_2: Eco RI;  
Constructed from size fractionated cytoplasmic RNA  
(4.4-7.4kb). Directionally cloned. Cells provided by  
Louis M. Staudt, Ph.D. Library preparation by Maria de  
Felima Bonaludo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT         35 a           90 C          53 g          52 t           1 others  
ORIGIN

Alignment Scores:

Pred. No.:	7.18e-23	Length:	231
Score:	34.00	Matches:	34
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.09%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-16 (1-1099) x AW501263 (1-231)

OY 994 leuSrAspnIllePheserArGlNserrSpvAlTrpsrPheglYvalleuTyR 1013  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 10 CTCCTGGACAACAATCTTCTCTGCCAGCAACAACGCTTGGAACCTTGCGGCTGCTCTGTAC 69

OY 1014 GluleuPherThrTYCYASpLySerCyScsrProserala 1027  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 70 GAGCTCTTACCTACTGCGAACAAAGGCGAGCCCTGCGGC 111

RESULT 38  
AM501582 254 bp mRNA linear EST 01-MAR-2000  
LOCUS  
DEFINITION UT-HF-BPDP-a-jg-b-03-0-UI.r1 NIH\_MGC\_51 Homo sapiens CDNA clone  
IMAGE:3074236 5', mRNA sequence.  
ACCESSION AM501582  
VERSION AM501582.1 GI:7115248  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>  
1 (bases 1 to 254)  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Eco RI site shown at the beginning of the sequence.  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: M.B. Soares Lab  
CDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bdrrp/image/image.html](http://www-bio.llnl.gov/bdrrp/image/image.html)  
Seq primer: M13 forward.

FEATURES  
Source Location/Qualifiers  
1..254  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3074236"  
/clone\_lib="NIH\_MGC\_51"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/lab\_host="DH10B (LTR)"  
/note="Vector: pT73-Pac; Site\_1: NotI; Site\_2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA  
(4.4-7.4kb). Directionally cloned. Cells provided by  
Louis M. Staudt, Ph.D. Library preparation by Maria de  
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 37 a 97 c 64 g 55 t 1 others  
ORIGIN

Alignment Scores:  
Pred. NO.: 8.04e-23 Length: 254  
Score: 34.00 Matches: 34  
Percent Similarity: 100.00% Conservative: 0  
Best local Similarity: 100.00% Mismatches: 0  
Query Match: 3.09% Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x AM501582 (1-254)

OY 994 LeuSerAspAsnIlePheSerArgInserAspValTTPSerPheGlyValValLeuTyr 1013  
|||||  
DB 10 CTCCTGGACAAACATCTTCTCTCGCCAGTCAGACGCTCGGGGCTGCTCTGTAC 69  
|||||

OY 1014 GluLeuPheThrTyrCysAspIysSerCysSerProSerAla 1027  
|||||  
DB 70 GAGCTCTTCACCTACTGCGACAAAGCTGACGCCCTCGGCC 111  
|||||

RESULT 39  
AM502902 286 bp mRNA linear EST 01-MAR-2000  
LOCUS  
DEFINITION UT-HF-BPDP-a-aiw-a-01-0-UI.r1 NIH\_MGC\_51 Homo sapiens CDNA clone  
IMAGE:3073056 5', mRNA sequence.  
ACCESSION AM502902  
VERSION AM502902.1 GI:7117812  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>  
1 (bases 1 to 286)  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Eco RI site shown at the beginning of the sequence.  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: M.B. Soares Lab  
CDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bdrrp/image/image.html](http://www-bio.llnl.gov/bdrrp/image/image.html)  
Seq primer: M13 forward.

FEATURES  
Source Location/Qualifiers  
1..286  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3073056"  
/clone\_lib="NIH\_MGC\_51"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/lab\_host="DH10B (LTR)"  
/note="Vector: pT73-Pac; Site\_1: NotI; Site\_2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA  
(4.4-7.4kb). Directionally cloned. Cells provided by  
Louis M. Staudt, Ph.D. Library preparation by Maria de  
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 43 a 110 c 68 g 65 t  
ORIGIN

Alignment Scores:  
Pred. NO.: 9.26e-23 Length: 286  
Score: 34.00 Matches: 34  
Percent Similarity: 100.00% Conservative: 0  
Best local Similarity: 100.00% Mismatches: 0  
Query Match: 3.09% Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x AM502902 (1-286)

OY 994 LeuSerAspAsnIlePheSerArgInserAspValTTPSerPheGlyValValLeuTyr 1013  
|||||  
DB 10 CTCCTGGACAAACATCTTCTCGCCAGTCAGACGCTCGGGGCTGCTCTGTAC 69  
|||||

OY 1014 GluLeuPheThrTyrCysAspIysSerCysSerProSerAla 1027  
|||||  
DB 70 GAGCTCTTCACCTACTGCGACAAAGCTGACGCCCTCGGCC 111  
|||||

RESULT 40  
AM501452 437 bp mRNA linear EST 01-MAR-2000  
LOCUS  
DEFINITION UT-HF-BPDP-a-jc-g-10-0-UI.r1 NIH\_MGC\_51 Homo sapiens CDNA clone  
IMAGE:3074106 5', mRNA sequence.  
ACCESSION AM501452  
VERSION AM501452.1 GI:7114998  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>  
1 (bases 1 to 437)  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Eco RI site shown at the beginning of the sequence.

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: M.B. Soares Lab  
CDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbp/image/image.html  
Seq primer: M13 forward.

FEATURES  
source  
1. 437  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="3074106"  
/clone\_lib="NIH\_MGC\_51"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/lab\_host="DH10B (LTI)"  
/note="Vector: pTZ19-Fac; Site\_1: NotI; Site\_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (4.4-7.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT  
63 a 180 c 105 g 89 t

ORIGIN  
Alignment Scores:  
Pred. No.: 1.53e-22 Length: 437  
Score: 34.00 Matches: 34  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.09% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AW501452 (1-437)

QY 994 LeuSerAspAsnIlePheSerArgInserAspValTrpSerPheGlyValValLeuTyr 1013  
|||||  
DB 10 CTCTCGACACATCTCTCTCCAGTCAGAGCTGTGGGGGCTGCTGTAC 69

QY 1014 GluLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027  
|||||  
DB 70 GAGCTCTTCACTACTGCGACAAAGCTGACGCCCTCGGCC 111

RESULT 41  
AM500984 451 bp mRNA linear EST 01-MAR-2000  
LOCUS  
DEFINITION  
IMAGE:3073676 5', mRNA sequence.  
ACCESSION  
AM500984  
VERSION  
AM500984.1 GI:7114087  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 451)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Eco RI site shown at the beginning of the sequence.  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: M.B. Soares Lab  
CDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbp/image/image.html  
Seq primer: M13 forward.  
Location/Qualifiers  
1. 451

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="3073676"  
/clone\_lib="NIH\_MGC\_51"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/lab\_host="DH10B (LTI)"  
/note="Vector: pTZ19-Fac; Site\_1: NotI; Site\_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (4.4-7.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT  
63 a 186 c 106 g 94 t 2 others

ORIGIN  
Alignment Scores:  
Pred. No.: 1.59e-22 Length: 451  
Score: 34.00 Matches: 34  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.09% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AW500984 (1-451)

QY 994 LeuSerAspAsnIlePheSerArgInserAspValTrpSerPheGlyValValLeuTyr 1013  
|||||  
DB 10 CTCTCGACACATCTCTCTCCAGTCAGAGCTGTGGGGGCTGCTGTAC 69

QY 1014 GluLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027  
|||||  
DB 70 GAGCTCTTCACTACTGCGACAAAGCTGACGCCCTCGGCC 111

RESULT 42  
AW408832 462 bp mRNA linear EST 16-FEB-2000  
LOCUS  
DEFINITION  
IMAGE:3063389 5', mRNA sequence.  
ACCESSION  
AW408832  
VERSION  
AW408832.1 GI:6927889  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 462)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Eco RI site shown at the beginning of the sequence.  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: M.B. Soares Lab  
CDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbp/image/image.html  
Seq primer: M13 forward.  
Location/Qualifiers  
1. 462

FEATURES  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="3063389"  
/clone\_lib="NIH\_MGC\_38"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/lab\_host="DH10B (LTI)"  
/note="Vector: pTZ19-Fac; Site\_1: NotI; Site\_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA

(2.5-3.5kb). Directionally cloned. Cells provided by Louis M. Staedt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 86 a 156 c 131 g 89 t

## ALIGNMENT SCORES:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1.64e-22	34.00	100.00%	100.00%	3.09%	462	34	0	0	0	0

US-09-397-967-16 (1-1099) x AM408832 (1-462)

QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957

DB 265 CGCGCTGCGTGCACGCGACCTGCGCGCGCGAACAATCTCTGAGAGCGAGCACAC 324

QY 958 ValIysIleAlaAspPheGlyIleuAlaIlysIleuProIeu 971

DB 325 GTCAAGATCGCTGACTTGGCTAGCTAAGCTGCTGCGCTT 366

RESULT 43 BF546077 478 bp mRNA linear EST 11-DEC-2000

LOCUS UI-R-C2P-rq-h-07-0-UI.r1 UI-R-C2P Rattus norvegicus cDNA clone

ACCESSION BF546077

VERSION BF546077.1 GI:11637184

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 478)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

Journal MEDLINE

COMMENT

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

CDNA Library Preparation: M.B. Soares Lab Clone distribution:

clones will be available through Research Genetics (www.resgen.com)

This clone is also available through the I.M.A.G.E. Consortium at

LNL (info@image.llnl.gov). IMAGE ID= 1790156

Seq primer: M13 Forward.

Location/Qualifiers

1. 478

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="UI-R-C2P-rq-h-07-0-UI"

/clone\_1id="UI-R-C2P"

/dev\_stage="adult"

/lab\_host="DHI0B (Life Technologies)"

/note="Vector: pUT3D-Pac (Pharmacia) with a modified

polylinker. Site\_1: Not I; Site\_2: Eco RI; The UI-R-C2P

library is a subtracted library derived from the UI-R-C1

library, which is a subtracted library derived from the

UI-R-C0 library. The UI-R-C0 library consisted of a

mixture of individually tagged normalized libraries

constructed from rat placenta, adult lung, brain, liver,

kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day

embryo. The tag is a string of 3-5 nucleotides present

between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2P) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DHI0B bacteria (Life Technologies) to generate the UI-R-C2P library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)"

BASE COUNT 83 a 167 c 128 g 99 t 1 others

## ALIGNMENT SCORES:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1.47e-19	31.00	100.00%	100.00%	2.82%	478	31	0	0	0	0

US-09-397-967-16 (1-1099) x BF546077 (1-478)

QY 988 TrpYrAlaProGluSerLeuSerAspAsnIlePheSerArgIleAspValTrpSer 1007

DB 9 TTGTAAGCCCTGAGTCTGATGACACATCTTCTCCGCAATGTGACGTGAGAC 68

QY 1008 PheGlyValIleLeuTYrGluLeuPheThrTYr 1018

DB 69 TTGCGCTGCTGCTGTATGACCTCTTACCTAC 101

## RESULT 44

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Soares, M.B., Rebelz, M., Pardinas, J., Liu, L. and Larson

W.M. Keck Center for Comparative and Functional Genomics

University of Illinois at Urbana-Champaign

340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL

61801, USA

Tel: 217 333 5998

Fax: 217 244 5617

Email: h-lewin@uiuc.edu

Research funding for cattle EST sequencing was provided by the USDA National

Research Initiative, Animal Genome Resource Grant AG 99-3205-6534

to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED

from Washington University Genome Center. Vector Trimmi g:

Cross-match from Washington University Genome Center PHRAP suite.

Sequences submitted are vector free and at least 200 bp in length.

PCR Primers

FORWARD: TAAATACGACTACTATAGGG

BACKWARD: ATATACGCTCCTACTAAG

Insert length: 483 Std Error: 0.00

Plate: BP230011A10 row: C column: 12

Seq primer: AGCGATATACATTTTCACACAGGA

FEATURES High quality sequence stop: 483.  
Location/Qualifiers

1. 483

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_1ib="BP23011A10C12"

/sex="female"

/lab\_host="DH10B"

/note="Organ: Placenta; Vector: pT7T3pac; Site\_1: ECORT; Site\_2: NOT; The cDNA library was contributed by the

Soares laboratory and it was constructed and normalized

as described by Bonaldo, M.F., Lennon, G. and Soares,

M.B. (1996), Genome Research 6(9): 791-806. "

BASE COUNT 83 a 185 c 121 g 94 t

ORIGIN

Alignment Scores:

Pred. No.: 1,42e-18

Score: 30.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2,73%

DB: 10

Gaps: 0

Length: 483

Matches: 30

Conservative: 0

Mismatches: 0

Indels: 0

US-09-397-967-16 (1-1099) x AM462965 (1-483)

RESULT 45

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

CONTACT

CELLULAR

HEINRICH-PETTE-INSTITUTE

MARTINSTR. 52, 20251 HAMBURG, GERMANY

EMAIL: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

FEATURES

source

1. /763

/organism="Gallus gallus"

/strain="CB"

/db\_xref="taxon:9031"

/clone="9d23r1"

/clone\_1ib="dkfz426"

/tissue\_type="Bursa of Fabricius"

/cell\_type="Bursal lymphocyte"

/dev\_stage="2-3 weeks old"

BASE COUNT 146 a 234 c 255 g 121 t 7 others

ORIGIN

Alignment Scores:

Pred. No.: 2,45e-18

Score: 30.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2,73%

DB: 9

Gaps: 0

Length: 763

Matches: 30

Conservative: 0

Mismatches: 0

Indels: 0

US-09-397-967-16 (1-1099) x AJ399251 (1-763)

RESULT 46

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

CELLULAR

HEINRICH-PETTE-INSTITUTE

MARTINSTR. 52, 20251 HAMBURG, GERMANY

EMAIL: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

FEATURES

source

1. /769

/organism="Gallus gallus"

/strain="CB"

/db\_xref="taxon:9031"

/clone="17j3r1"

/clone\_1ib="dkfz426"

/tissue\_type="Bursa of Fabricius"

/cell\_type="Bursal lymphocyte"

/dev\_stage="2-3 weeks old"

BASE COUNT 150 a 222 c 262 g 134 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 2,47e-18

Score: 30.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2,73%

DB: 9

Gaps: 0

Length: 769

Matches: 30

Conservative: 0

Mismatches: 0

Indels: 0

US-09-397-967-16 (1-1099) x AJ442008 (1-769)

RESULT 47

LOCUS

DEFINITION

UT-HF-BNO-akv-c-08-0-UI.r1 NIH\_MGC\_50 Homo sapiens cDNA clone

AM503697

411 bp mRNA linear EST 01-MAR-2000

IMAGE:3078518 5', mRNA sequence.  
 ACCESSION AM503697.1 GI:7140506  
 VERSION EST.  
 KEYWORDS human.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 411)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Eco RI site shown at the beginning of the sequence.  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 CDNA Library Preparation: M.B. Soares Lab  
 CDNA Library Arrayed by: M.B. Soares Lab  
 DNA Sequencing by: M.B. Soares Lab  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html  
 Seq primer: M13 Forward

FEATURES  
 source  
 1..411  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3078518"  
 /clone\_lib="NIH-MGC-50"  
 /tissue\_type="lymph"  
 /cell\_type="germinal center B cells"  
 /lab\_host="DH10B (LTR)"  
 /note="Vector: pUTR-Pac; Site\_1: NotI; Site\_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (3.5-4.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 69 a 131 c 141 g 70 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.07e-16 Length: 411  
 Score: 28.00 Matches: 28  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.55% Indels: 0  
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AM503697 (1-411)

QY 173 MetalaatgglglnalaglAarproglgltuleuLeuLysThrValserTyrlysAla 192  
 |||||||  
 DB 193 ATGGCGGAGAGACAGCCAGCGCGGAGAGAGCTGCTGAGACGTCTACGACACAGGCC 252  
 |||||||

QY 193 CysleuPProPtoSerLeuAarasp 200  
 |||||||  
 DB 253 TGCCATACCCCAAGCCTCGCGAC 276  
 |||||||

RESULT 48  
 BE441858 491 bp mRNA linear EST 01-DEC-2000  
 LOCUS BE441858  
 DEFINITION 258301 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION BE441858  
 VERSION BE441858.1 GI:11501950  
 KEYWORDS EST.  
 SOURCE pig.  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 REFERENCE 1 (bases 1 to 491)  
 Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E., Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.

TITLE and Keele, J.W.  
 DESIGN and use of two pooled tissue normalized cDNA libraries for  
 EST discovery in swine  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@emall.marc.usda.gov  
 Single pass sequencing. Bases called and alt-trimmed with phred  
 v0.980904.e. Vector identified by cross-match with the -m1nscore 18  
 and -mismatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCTATGACAT  
 BACKWARD: GTTTCACGACGACGACG  
 Plate: 87 row: M column: 9  
 Seq primer: ATTAGTGACCTATAG.

FEATURES  
 source  
 1..491  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /clone\_lib="MARC 2P1G"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 100 a 155 c 134 g 102 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.25e-15 Length: 491  
 Score: 27.00 Matches: 27  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.46% Indels: 0  
 DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BE441858 (1-491)

QY 668 GlyValserProThrValleuserLeuGluMetleuThrAspArgIleProTrpValAla 687  
 |||||||  
 DB 3 GGTGTACGCCACCTGTACTAGCTGAGATGCTTACGACAGATCCCGGGGGGCC 62  
 |||||||

QY 688 ProGluCysLeuGlnGluAla 694  
 |||||||  
 DB 63 CCTGAGTGTCTCCAGAGGCC 83  
 |||||||

RESULT 49  
 BE309453 550 bp mRNA linear EST 26-OCT-2000  
 LOCUS BE309453  
 DEFINITION 601095351P1 NCL\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:3489805 5',  
 mRNA sequence.  
 ACCESSION BE309453  
 VERSION BE309453.1 GI:9167666  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 550)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:



http://image.jnl.gov  
Plate: L14M8531 row: f column: 14  
High quality sequence stop: 532.  
Location/Qualifiers  
1.550

FEATURES  
source

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3489805"  
/clone\_1lb="NCI\_CGAP\_Mam5"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;  
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH"

BASE COUNT 97 a 153 c 158 g 142 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1.43e-15 Length: 550  
Score: 27.00 Matches: 27  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.46% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x BE309453 (1-550)

OY 111 G1yleuglunrCyHsIarGpHeG1yleuArGysAspleuThrSerAlaIleuAsp 130  
|||||  
DB 369 GGGCTGGAGACATGTCACCGCTTGGCTGGCGACAGATTGACCACTGCATCTTGAC 428  
|||||  
OY 131 LeuHsValleuglunHsleu 137  
|||||  
DB 429 TTACATGTATTGAAACATCTC 449  
|||||

RESULT 50  
BF191620 526 bp mRNA linear EST 02-NOV-2000  
LOCUS 239255 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
DEFINITION BF191620  
ACCESSION  
VERSION BF191620.1 GI:11074989  
KEYWORDS EST.  
SOURCE  
ORGANISM Sus scrofa

REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
Mammalia: Eutheria: Cetartiodactyla: Suina: Suidae: Sus.  
1 (bases 1 to 526)  
AUTHORS Fahnenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,  
Stone,R.T., Heaton,M.P., Grose,W.M., Bennett,G.A., Laegreid,W.W.,  
and Keele,J.W.

TITLE Design and use of two pooled tissue normalized cDNA libraries for  
EST discovery in swine  
JOURNAL Unpublished (2000)  
COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt\_trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -mnscore 18  
and -minmatch 12 options.

FEATURES  
source

PCR primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCAGTCACGACG

Plate: 67 row: D column: 8  
Seq primer: ATTACGTGACACTATG.  
Location/Qualifiers  
1.526

/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_1lb="MARC 2P1G"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
Library made from pooled tissue from testis, ovary,  
endometrium, hypothalamus, pituitary, and placenta."

## Alignment Scores:

Pred. No.: 1.24e-13 Length: 526  
Score: 25.00 Matches: 25  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.27% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BF191620 (1-526)

OY 860 G1nglNarGAspHeG1nArG1nleG1nleuYsAlaIleuHsSerAspHele 879  
|||||  
DB 432 CAGCAGAGGACTTCCAGCGGAGATCCAGATCTTAAGCCCTCCACAGCCTCAT 491  
|||||  
OY 880 ValLySTyTArGly 884  
|||||  
DB 492 GTCAAGTACCGGGG 506  
|||||

RESULT 51  
BM088287 564 bp mRNA linear EST 19-NOV-2001  
LOCUS 501721 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.  
DEFINITION BM088287  
ACCESSION BM088287.1 GI:16998915  
VERSION  
KEYWORDS EST.  
SOURCE  
ORGANISM Bos taurus

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
Mammalia: Eutheria: Cetartiodactyla: Ruminantia: Pecora: Bovidae;  
Bovidae: Bovinae; Bos.  
1 (bases 1 to 564)  
AUTHORS Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
Casas,E., Wray,J.E., White,J., Cho,J., Fahnenkrug,S.C., Bennett,  
G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,  
Petra,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and  
Keele,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
JOURNAL Genome Res. 11 (4), 626-630 (2001)  
MEDLINE 21180013  
COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt\_trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -mnscore 18  
and -minmatch 12 options.

FEATURES  
source

PCR primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCAGTCACGACG

Plate: 140 row: C column: 18  
Seq primer: ATTACGTGACACTATG.  
Location/Qualifiers  
1.564

/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_1lb="MARC 2BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;

Library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."

BASE COUNT 116 a 176 c 161 g 111 t

## ALIGNMENT SCORES:

Pred. No.: 1.34e-13 Length: 564  
Score: 25.00 Matches: 25  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.27% Indels: 0  
Caps: 0

US-09-397-967-16 (1-1099) x BM088287 (1-564)

QY 670 SerProthValleuSerleuGlumelLeuthrAsparGileProtpValAlaProGlu 669  
|||||

DB 3 AGCCCACTGTGCTAGAGCTGAGATGCTCAGTACAGAGATCCCTGGTGGCCCTGAA 62

QY 690 CysleuGlnGluAla 694  
|||||

DB 63 TGTCTCCAGAGAGGCC 77

## RESULT 52

LOCUS B0709825 871 bp mRNA linear EST 16-JUL-2002  
DEFINITION AGEMCOURT.8474895 NIH\_MGC\_113 Homo sapiens CDNA clone IMAGE:6301440

ACCESSION B0709825  
VERSION B0709825.1 GI:21848724

KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 871)  
NIH-MGC <http://mgc.ncl.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Dr. Mark Watson

CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LICM2517 row: f column: 01  
High quality sequence stop: 595.

FEATURES  
Location/Qualifiers

1..871  
/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:6301440"

/clone\_id="NIH\_MGC\_113"

/lab\_host="DH10B (phage-resistant)"

/note="Organ made by oligo-df priming. Directionally cloned  
ECORI; CDNA made by oligo-df priming. Directionally cloned  
into ECORI/XhoI sites using the following 5' adaptor:  
GGGACGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH-MGC Library."

BASE COUNT 149 a 300 c 264 g 158 t

ORIGIN  
Alignment Scores: 2.25e-13 Length: 871  
Pred. No.: 25.00 Matches: 25  
Score: 100.00% Conservative: 0  
Percent Similarity:

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.27% Indels: 0  
DB: 14 Caps: 0

US-09-397-967-16 (1-1099) x B0709825 (1-871)

QY 323 lLeleuGluAlaGluPheProGlyLeuProGlyAlaLeuSerPheValAlaLeuValasp 342  
|||||

DB 328 ATTTAGAGCGCGAGTTCACAGGCGCTGCCAGAGCTGTGTCGTCGCGGCGCTGGAGC 387

QY 343 GlyTyrPheArgLeu 347  
|||||

DB 388 GGCTACTTCCGGCTG 402

## RESULT 53

LOCUS BR791980 371 bp mRNA linear EST 16-NOV-2001  
DEFINITION BR791980 RIKEN full-length enriched, brain CRT-1443 BC3H1 CDNA Mus  
musculus CDNA clone GA30139C08 3, mRNA sequence.

ACCESSION BR791980  
VERSION BR791980.1 GI:16960744

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 371)  
Akimura,T., Arahata,T., Carninci,P., Furuno,M., Hanagaki,T., Ishii

Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii

Y., Ito,M., Kawai,D., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,

Nakamura,M., Nishii,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,

Saito,K., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,

Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa

,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toyota,T.,

Watanahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.,

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.

2001)

JOURNAL Unpublished (2001)

COMMENT Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC) Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222

Fax: 81-45-503-9216  
Email: [genome-res@sc.riken.go.jp](mailto:genome-res@sc.riken.go.jp),  
[URL:http://genome.gsc.riken.go.jp/](mailto:URL:http://genome.gsc.riken.go.jp/)

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh

,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.,

Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagii,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
Watanahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura

,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and  
Hayashizaki,Y.,

RIKEN integrated sequence analysis (RISA) system-384 format  
sequencing pipeline with 384 multipipillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara  
,Y. and Hayashizaki,Y.,  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
e mouse tissues.  
Location/Qualifiers  
1..371  
/organism="Mus musculus"  
/strain="C3H"  
/db\_xref="taxon:10090"  
/clone="G430139C08"

```

/dev_stage="10 months, virgin"
/lab_host="DH108"
/note="Organ: mammary; Vector: pCMV-Sport6; site_1: SalI;
site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      85 a      113 c      79 g      61 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:      6,66e-12      Length:      339
Score:      23.00      Matches:      23
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      2.09%      Indels:      0
DB:      10      Gaps:      0

US-09-397-967-16 (1-1099) x AM230832 (1-339)
Qy      1053      ArgArgLeuProPProPProPProhKcYsPProhGluValGlnGluLeuMetGlnLeuGys      1072
|||||
Db      3      CGAGCGCTCCACACACCTCCACCTCCACCGAGGTTGAGAGCTCATGACGCTGTC      62
|||||
Qy      1073      TrpAlaPro      1075
|||||
Db      63      TGGGCGGCC      71

RESULT 55
AM505513      540 bp      mRNA      linear      EST 02-MAR-20000
LOCUS      UI-HF-BN0-amb-g-09-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone
ACCESSION      IMAGE:3081784 5', mRNA sequence.
VERSION      AM505513
KEYWORDS      AM505513.1 GI:7143180
SOURCE      EST.
ORGANISM      human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 540)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab
cDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrp/image/image.html
Seq primer: M13 Forward.
Location/Qualifiers
1..540
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3081784"
/clone_1lb="NIH_MGC_50"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH108 (LTI)"
/note="Vector: pTrf3-Pac; Site_1: NotI; site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT      116 a      164 c      151 g      108 t      1 others
ORIGIN

```

Pred. No.: 1.1e-10 Length: 540  
 Score: 22.00 Matches: 22  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.00% Indels: 0  
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AW505513 (1-540)

QY 824 LeuGlyfysglAsnPhleglySerValGluLeuCyArGfYrAsProLeuGlyAspAsn 843  
 |||||||  
 DB 365 CTGGCAAGGCGACACTTGGCAGCGAGCTGTGCGCTATGACCGCTAGCGACAAI 424

OY 844 ThrGly 845  
 |||||||  
 DB 425 ACAGCT 430

RESULT 56

LOCUS AW785501 472 bp mRNA linear EST 09-JUL-2000  
 DEFINITION J16874 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION AW785501 GI:7842349

VERSION AW785501.1  
 KEYWORDS EST.  
 SOURCE pig.  
 ORGANISM Sus scrofa

REFERENCE 1 (bases 1 to 472)  
 AUTHORS Fahrner, S.C., Freking, B.A., Rohrer, G.A., Shultz, P.T., Casas, E.,  
 Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,  
 and Keefe, J.W.  
 Design and use of two pooled tissue normalized cDNA libraries for  
 EST discovery in swine  
 Unpublished (2000)

JOURNAL COMMENT

USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390

Email: smithem@marc.usda.gov  
 Simple pass sequencing. Bases called and alt-trimmed with phred  
 v0.980904.e. Vector identified by cross-match with the -mmscore 18  
 and -mismatch 12 options.

PCR Primers  
 FORWARD: AGGAACAGCATGACCAT  
 BACKWARD: GTTTCACGACGACGACG  
 Plate: 52 row: D column: 2  
 Seq primer: ATTGAGTGACCATATAG.

FEATURES

source 1. 472  
 Location/Qualifiers

/organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /clone\_lib="MARC 1P1G"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV SPORT6; Site\_1: NCI; Site\_2: Sali;  
 library made from pooled tissue from day 11, 13, 15, 20,  
 and 30 embryos."

BASE COUNT 117 a 128 c 134 g 93 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 8.95e-10 Length: 472  
 Score: 21.00 Matches: 21  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.91% Indels: 0  
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AW785501 (1-472)

OY 578 HisLeuValLeuLeuHisGlyValAlcysMetAlaGlyAspSerIleMetValGlnGluPhe 597

DB 322 CATCTGCTTCTCCATGCGCTATGCATGCTGAGACAGCATCATGTCGAGAAATTC 381  
 |||||||  
 OY 598 Val 598  
 |||||  
 DB 382 GTA 384

RESULT 57

LOCUS BB590115 258 bp mRNA linear EST 30-NOV-2000  
 DEFINITION BB590115 RIKEN full-length enriched, 0 day neonate thymus Mus  
 musculus cDNA clone AA30103J21 5', mRNA sequence.  
 BB590115

ACCESSION BB590115 GI:11486659  
 VERSION  
 KEYWORDS EST.

SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 258)  
 AUTHORS Alizawa, K., Akehira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P.,  
 Hanagaki, T., Hayatsu, N., Hiraoka, T., Hirozane, T., Hodojima, Y.,  
 Imocani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J., Kojima, Y., Konno,  
 H., Kusakabe, M., Matsuyama, T., Miyazaki, A., Nakamura, M., Nishi, K.,  
 Nomura, K., Numazaki, K., Okazaki, Y., Okido, T., Owa, C., Sakai, C.,  
 Sakai, K., Sasaki, D., Sato, K., Shibata, K., Shibata, Y., Shingawa, A.,  
 Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka,  
 T., Toyota, T., Watahiki, A., Yamamura, T., Yasunishi, A., Yoshida, K.,  
 Yoshiki, A., Muramatsu, M., and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Alizawa, K. et al. 2000)  
 Unpublished (2000)

JOURNAL COMMENT  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

Email: genome-resgsc.riken.go.jp/  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki,  
 N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 Thermostabilization and thermoactivation of thermolabile enzymes by  
 trehalose and its application for the synthesis of full length  
 cDNA. Proc Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh, M., Kitsuina, T., Akiyama, J., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
 Y., and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation  
 system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
 19-44 (1999)

Please visit our web site (http://genome.rtc.riken.go.jp) for  
 further details.

FEATURES

source

Location/Qualifiers

1. 258  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone\_lib="AA30103J21"  
 /clone\_lib="RIKEN full-length enriched, 0 day neonate  
 thymus"  
 /tissue\_type="thymus"  
 /dev\_stage="0 day neonate"  
 /lab\_host="DH10B"  
 /note="Site\_1: Sali; Site\_2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTTTVN 3'], cDNA was

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lib="IMAGE:2304105"
/clone_1lib="NCI_CGAP_GC6"
/tissue_type="pooled germ cell tumors"
/lab_host=DH10B
/notes=Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP_GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonids
1257096-1258631, 1469004-1470983, and 1475592-1476743).

```

Alignment Scores:	
Pred. No.:	4,866-07
Score:	18.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Length:	319
Matches:	18
Conservative:	0
Mismatches:	0

Query Match: 1.64% Indels: 0  
DB: 13 Gaps: 0

US-09-397-967-16 (1-1099) x BM032194 (1-319)

QY 559 AsncysmetcIusRphelugluAlaIalaserLeumetserGlnValserTyr 576  
|||||  
DB 112 AATTGCATGAGTCATTCCTCGAAGACGAGCTTGATGACCAAGTCTCATAC 59

RESULT 60  
BM146700 373 bp mRNA linear EST 30-NOV-2001  
LOCUS TCAAP1E7063 pediatric acute myelogenous leukemia cell (FAB M1)  
DEFINITION Baylor-HSC project-TCAA Homo sapiens CDNA clone TCAAP1063, mRNA  
sequence.  
ACCESSION BM146700  
VERSION BM146700.1 GI:17165865  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 373)  
AUTHORS Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr.,  
Gunsarane, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.  
TITLE Pediatric leukemia CDNA Sequencing Project (2001)  
JOURNAL Unpublished (2001)  
COMMENT Contact: Dr. Judith F. Margolin  
Texas Children's Cancer Center and Human Genome Sequencing Center  
at Baylor College of Medicine  
1102 Bates, MC3-3320 Houston, TX 77030, USA  
Tel: 832-824-4536  
Fax: 832-825-4038  
Email: clones@ccc.org  
Seq primer: M13 primer.  
Location/Qualifiers  
1. 373  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="TCAAP1063"  
/clone\_lib="pediatric acute myelogenous leukemia cell (FAB M1) Baylor-HSC project-TCAA"  
/sex="male"  
/tissue\_type="leukopheresis"  
/cell\_type="myeloid cell"  
/dex\_stage="pediatric 6 years"  
/lab\_host="DH10B"  
/note="Vector: lambda PSB; Site\_1: BamHI; Site\_2: EcoRI;  
First strand cDNA was primed with an anchored  
XhoI-oligo(dT) primer [5'GGAGGACTCGACGCCGCCGAGAGAG(T)VN  
3'; V-A,C,G; N-A,C,G,T] and then dg tailed. Second strand  
was primed with a BamHI-dc primer  
[5'AGAGAGCTCGATCCGCCGCCGATATATATAT(C) 3'].  
Double-stranded cDNA was then digested with BamHI and XhoI  
and directionally cloned into the BamHI and SalI sites of  
lambda PSB vector. Library went through one round of  
normalization. Library was constructed by Wei Yu at RIKEN  
of Japan (Carninci P., Westover A., Nishiyama Y., Ohsumi T.,  
Itoh M., Nagaoka S., Sasaki, Y., Muramatsu M.,  
Schneider C., Hayashizaki Y., High efficiency selection of  
full-length cDNA by improved biotinylated cap trapper.,  
DNA Res 4: 1, 61-6, Feb 28, 1997)".

BASE COUNT 51 a 138 c 101 g 83 t  
ORIGIN

Alignment Scores: 5.85e-07 Length: 373  
Pred. No.: 18.00 Matches: 18  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 1.64% Gaps: 0  
DB: 13

US-09-397-967-16 (1-1099) x BM146700 (1-373)

QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeu 18  
|||||  
DB 63 ATGGACCTCCAAAGTGAAGACAGCCCTTGATCCCTCAGCGTTGATGACGCTC 116

RESULT 61  
AM504672 413 bp mRNA linear EST 202-MAR-2000  
LOCUS UT-HF-BNO-all-d-09-0-UT.r1 NIH-MGC\_50 Homo sapiens CDNA clone  
DEFINITION IMAGE:3080104 5', mRNA sequence.  
ACCESSION AM504672  
VERSION AM504672.1 GI:7142339  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 413)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Eco RI site shown at the beginning of the sequence.  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: M.B. Soares Lab  
CDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/bdrrp/image/image.html  
The following repetitive elements were found in this cDNA  
sequence:  
121-152. >GC rich flow complexity  
Seq primer: M13 forward  
Location/Qualifiers  
1. 413  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3080104"  
/clone\_lib="NIH-MGC\_50"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/lab\_host="DH10B (LTR)"  
/note="Vector: p7713-Pac; Site\_1: NotI; Site\_2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA  
(3.5-4.4kb). Directionally cloned. Cells provided by  
Louis M. Staudt, Ph.D. Library preparation by Maria de  
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 63 a 142 c 114 g 94 t  
ORIGIN

Alignment Scores: 6.61e-07 Length: 413  
Pred. No.: 18.00 Matches: 18  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 1.64% Gaps: 0  
DB: 10

US-09-397-967-16 (1-1099) x AM504672 (1-413)

QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeu 18  
|||||  
DB 32 ATGGACCTCCAAAGTGAAGACAGCCCTTGATCCCTCAGCGTTGATGACGCTC 85

RESULT 62  
AO192794 481 bp DNA linear GSS 04-NOV-1998  
LOCUS HS\_2251\_B2\_P07\_MR CIT Approved Human Genomic Sperm Library D Homo  
DEFINITION sapiens genomic clone Plate-2251 Col-14 Row=L, DNA sequence.

ACCESSION AQ192794  
 VERSION AQ192794.1 GI:3591416  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 481)  
 AUTHORS Mahalax,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
 Hood L.  
 TITLE Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
 MEDLINE 99380589  
 COMMENT Contact: Mahalax GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Sequence Tagged Connector  
 Plate: 2251 row: L column: 14  
 Class: BAC ends  
 High quality sequence stop: 481.  
 Location/Qualifiers  
 1..481  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_plate="2251 Col-14 Row-L"  
 /clone\_lib="CIT Approved Human Genomic Sperm Library D"  
 /sex="male"  
 /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in  
 E-Coli DH10B"  
 BASE COUNT 101 a 115 c 192 g 72 t 1 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.92e-07 Length: 481  
 Score: 18.00 Matches: 18  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.64% Indels: 0  
 DB: 17 Gaps: 0  
 US-09-397-967-16 (1-1099) x AQ192794 (1-481)  
 QY 943 ATGAGPLEUAlAAlaArgSntleuVaJGUsenGluAlaHisVallystle 960  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 432 CCGGACCTGGCCGCCGGAACAATCCTCGTGAGAGCGAGCACACCTCAAGATC 379  
 RESULT 63 781 bp mRNA linear EST 16-OCT-2001  
 BI906136  
 LOCUS 603062556r1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5211901 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI906136  
 VERSION BI906136.1 GI:16168788  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 781)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)

DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ULNL at:  
 http://image.llnl.gov  
 Plate: LLM11531 row: p column: 14  
 High quality sequence stop: 581.  
 Location/Qualifiers  
 1..781  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="5211901"  
 /clone\_lib="NIH\_MGC\_118"  
 /tissue\_type="Leukocyte"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6; Site1: NotI; Site2: EcoRV  
 (destroyed); RNA source leukocytes from anonymous pool of  
 non-activated adult donors. Library is oligo-dT primed  
 and directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.7 kb, insert size range  
 1.2-3.3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 027. Note:  
 this is a NIH\_MGC Library."  
 BASE COUNT 242 a 207 c 184 g 148 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.41e-06 Length: 781  
 Score: 18.00 Matches: 18  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.64% Indels: 0  
 DB: 13 Gaps: 0  
 US-09-397-967-16 (1-1099) x BI906136 (1-781)  
 QY 559 AsncysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyr 576  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 480 AACTGCATGAGTCATCTCTGGAGAGCGACCTGATGAGCCAACTGTCGTAC 533  
 RESULT 64 814 bp mRNA linear EST 16-OCT-2001  
 BI911448  
 LOCUS 603063362r1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5212575 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI911448  
 VERSION BI911448.1 GI:16175189  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 814)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ULNL at:  
 http://image.llnl.gov  
 Plate: LLM11533 row: 1 column: 16  
 High quality sequence stop: 5  
 High quality sequence stop: 416.  
 Location/Qualifiers  
 1..814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="5212575"

```
/clone_id="NIH_MGC_118"  
/tissue_type="leukocyte"  
/lab_host="DH10B"  
/note="Vector: pCMV-SPORT6, Site_1: NotI, Site_2: EcoRV  
(destroyed); RNA source leukocytes from anonymous pool of  
non-activated adult donors. Library is oligo-dT primed  
and directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 kb, insert size range  
1.2-3.3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 027. Note:  
this is a NIH_MGC library."  
BASE COUNT      168 a      258 c      237 g      151 t  
ORIGIN  
  
Alignment Scores:  
Pred. NO.:      1.48e-06      Length:      814  
Score:          18.00      Matches:      18  
Percent Similarity: 100.00%      Conservative: 0  
Best Local Similarity: 100.00%      Mismatches: 0  
Query Match:    1.64%      Indels:      0  
DB:             13      Gaps:          0  
  
US-09-397-967-16 (1-1099) x B1911448 (1-814)  
  
QY      1      MetaIaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeu 18  
|||||  
B147824/c 90      ATGCACCTCCAGAGTGAAGACAGCCCTGATCCCTCATGCTATATGACGCTC 143  
|||||  
  
RESULT 65  
B147824/c 313 bp      mRNA      linear      EST 29-DEC-2000  
LOCUS      B147824      313 bp      mRNA      linear      EST 29-DEC-2000  
DEFINITION      uy83e09.x1 NCI_CGAP Mam5 Mus musculus cDNA clone IMAGE:366184 3'  
Similar to SW:JAK3_RAT Q63272 TYROSINE-PROTEIN KINASE JAK3 ;, mRNA  
sequence:  
B147824      B147824.1      GI:11029219  
EST.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 313)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Other_ESTs: uy83e09.y1  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
Image.lnl.gov/image/html/lresources.shtml  
MGI:1426952  
Seq primer: -40UP from GIBCO  
High quality sequence stop: 280.  
Location/Qualifiers  
1. 313  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db_xref="taxon:10090"  
/clone="IMAGE:366184"  
/clone_id="NCI_CGAP_Mam5"  
/tissue_type="tumor, gross tissue"  
/dev_stage="7 months"  
/lab_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6, Site_1: SalI;  
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
```

```
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Robin Humphreys,  
NIH  
BASE COUNT      61 a      81 c      111 g      55 t      5 others  
ORIGIN  
  
Alignment Scores:  
Pred. NO.:      4.53e-06      Length:      313  
Score:          17.00      Matches:      17  
Percent Similarity: 100.00%      Conservative: 0  
Best Local Similarity: 100.00%      Mismatches: 0  
Query Match:    1.55%      Indels:      0  
DB:             12      Gaps:          0  
  
US-09-397-967-16 (1-1099) x B147824 (1-313)  
  
QY      1059      ProthCysProThrGluValGlnGluLeuMetGlnCysTTPAlaPro 1075  
|||||  
DB      286      CCCACCTCCACCCACGAGGAGGAGCTCATGCGAGTGTCTGGCCGCC 236  
|||||  
  
RESULT 66  
BG995133/c 426 bp      mRNA      linear      EST 13-JUN-2001  
LOCUS      BG995133      426 bp      mRNA      linear      EST 13-JUN-2001  
DEFINITION      CM0-HT1295-160201-779-906 HT1295 Homo sapiens cDNA, mRNA sequence.  
ACCESSION      BG995133  
VERSION      BG995133.1      GI:14399203  
KEYWORDS      EST.  
SOURCE      human.  
ORGANISM      Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 426)  
REFERENCE  
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordalo,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
  
JOURNAL  
MEDLINE      20202663  
COMMENT      Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM06t2-CM0-HT1295-  
160201-779-906&t3=2001-02-16&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 10  
High quality sequence stop: 426.  
Location/Qualifiers  
1. 426  
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone_id="HT1295"  
/dev_stage="Adult"  
/note="Organ: head,neck; Vector: puc18; Site_1: SmaI;  
Site_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
BASE COUNT      82 a      120 c      122 g      102 t  
ORIGIN
```



Alignment Scores:

Pred. No.:	6.53e-06	Length:	426
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.55%	Indels:	0
DB:	13	Gaps:	0

US-09-397-967-16 (1-1099) x BG995133 (1-426)

QY 514 A9PSeuclutPrphsGuaneuGlyHicGlySerPherhrlYsile 530  
 DB 187 GACAGCCTGAGTGCATGAGACCTGGCCATGGCTCTTACCAAGATT 137

RESULT 67 430 bp mRNA linear EST 09-MAR-1999  
 A1472755 talle03.x1 NCI CGAP\_Lym5 Homo sapiens cDNA clone IMAGE:2043964 3'  
 LOCUS similar to SW:JAK3\_HUMAN P52333 TYROSINE-PROTEIN KINASE JAK3 ;  
 DEFINITION mRNA sequence.

ACCESSION A1472755  
 VERSION A1472755.1 GI:4325800  
 KEYWORDS EST.

SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 430)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@b-remail.nih.gov  
 Tissue Procurement: Mark Raffeld, M.D.  
 cDNA Library Preparation: Stratagene, Inc.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/MLM at:  
 www.bio.lnl.gov/bdrp/image/image.html

Trace considered overall poor quality  
 Seg primer: -40UP from Glibco  
 High quality sequence stop: 1.  
 Location/Qualifiers

1. .430  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2043964"  
 /clone\_lib="NCI-CGAP\_Lym5"  
 /issue\_type="follicular lymphoma"  
 /lab\_host="SOLR (Stratagene, karmaycin resistant)"  
 /note="Organ: lymph node; Vector: pBluescript SK-; Site\_1:  
 EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
 Oligo dt. Average insert size 1.2 kb. Non-amplified  
 library. -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3'  
 adaptor sequence: 5' CTCGACTTTTCTTTTCTTTT 3'"  
 BASE COUNT 95 a 135 c 97 g 103 t  
 ORIGIN

Alignment Scores:

Pred. No.:	6.6e-06	Length:	430
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.55%	Indels:	0
DB:	9	Gaps:	0

US-09-397-967-16 (1-1099) x A1472755 (1-430)

QY 560 CyswctGUSErPhelGUAAlaIsSerLeuMetSerGlnAlSerTyr 576  
 ||||||||||||||||||||||||||||||||||||||||||||

DB 94 TGCATGAGTCAATCTCTGAGACGACGAGCTTGATGACCAAGTGTCTAC 44

RESULT 68 449 bp mRNA linear EST 21-APR-1999  
 A1598913 EST250616 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone  
 LOCUS RENDY61.3' end, mRNA sequence.  
 DEFINITION

ACCESSION A1598913  
 VERSION A1598913.1 GI:4607961  
 KEYWORDS EST.  
 SOURCE Rattus sp.  
 ORGANISM Rattus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 449)  
 Lee,N.H., Gloder,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
 Kerlavage,A.R. and Adams,M.D.  
 Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat  
 Gene Index

JOURNAL Unpublished (1998)  
 COMMENT Other\_ESTs: TC82898  
 Contact: Lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@ligr.org  
 Seq primer: M13-21.

FEATURES  
 source Location/Qualifiers

1. .449  
 /organism="Rattus sp."  
 /db\_xref="taxon:10118"  
 /clone="RENDY61"  
 /clone\_lib="Normalized rat embryo, Bento Soares"  
 /dev\_stage="embryo 8, 12, 18 dpc"  
 /note="Vector: pT713Pac; Site\_1: EcoRI; Site\_2: NotI"  
 BASE COUNT 102 a 118 c 135 g 94 t  
 ORIGIN

Alignment Scores:

Pred. No.:	6.95e-06	Length:	449
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.55%	Indels:	0
DB:	9	Gaps:	0

US-09-397-967-16 (1-1099) x A1598913 (1-449)

QY 159 GlnGlyGUPheUSeSerLeuAlaValLeuAspLeuAlaGlnMetAlaarg 175  
 ||||||||||||||||||||||||||||||||||||||||||||

DB 420 CAGGACAGAGTCCGACGACCTAGCTGTCTGACCTGCCACAGATGGCTCGT 370  
 RESULT 69  
 BB285471  
 LOCUS BB285471 239 bp mRNA linear EST 09-JUL-2000  
 DEFINITION BB285471 RIKEN full-length enriched 2 cells egg Mus musculus cDNA  
 clone B020006C2.3', similar to I1656 Mouse protein-tyrosine kinase  
 (JAK2) mRNA, mRNA sequence.

ACCESSION BB285471  
 VERSION BB285471.1 GI:8985920  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 239)  
 Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carrincl  
 P., Endo,T., Fukuda,S., Fukumishi,Y., Hara,A., Hayatsu,N.,  
 Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,  
 Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,O., Kikuchi,N.,  
 Kiyosawa,H., Kojima,Y., Kondo,S., Koye,S., Kurihara,C., Kusakabe,M.

FEATURES	Location/Qualifiers
source	1. . 239

Alignment Scores:	
Pred. No.:	3,13e-05
Score:	16.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	1.46%
DB:	10
Length:	239
Matches:	16
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

TITLE	Shotgun sequencing of the human transcriptome with ORF sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000).
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G. Department of Genetics University of Cambridge

FEATURES	Location/Qualifiers
source	1. . 258

US-09-597-96/16 (1.1059) x 2100/204 (1.220)	
QY 1003 SetaspValIrrpserPieglyValValleuYrgIuIsePheThTyr 1018	
DB 154 TCAGATGCTTGACCTTGGAGTGGTCTGATGAGACCTTTTCACATAC 107	
RESULT 71	
BR285470	

RESULT 71	LOCUS	DEFINITION
BB285470	BB285470	304 bp mRNA linear EST 09-JUL-2000
	BB285470	RIKEN full-length enriched, 2 cells egg Mus musculus cDNA
	clone B020006c21	3' similar to I16956 Mouse protein-tyrosine kinase
	(Jnk2) mRNA, mRNA sequence	

BASE COUNT

Prod. NO.:	5.74e-05	Length:	398
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0

DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x AA453345 (1-398)

QY 1003 SeraspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
|||||  
DB 207 TCAGATGTTTGAGCTTGGAGTGTGTGTATGAACCTTTCACATAC 254

RESULT 73  
T48592 406 bp mRNA linear EST 02-FEB-1995  
LOCUS ph6h6\_19/1TV Outward Alu-primed hncDNA library Homo sapiens cDNA  
DEFINITION clone ph6h6\_19/1TV, mRNA sequence.  
T48592  
T48592.1 GI:642792  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Vinogradova,T.V., Lebedev,Y.B., Kopantzev,E.P., Wagner,L.L., Volik  
,S.V., Ermolaeva,O.D., Lavrentyeva,T., Monastyrskaya,G.S. and  
Sverdlov,E.D.  
TITLE Outward Alu-primed hncDNA library  
JOURNAL Unpublished (1995)  
COMMENT Contact: Sverdlov ED  
Structure and Function of Human Genes  
Shemyakin Institute of Bioorganic Chemistry  
16/10 Miklukho-Maklaya, Moscow, 117871, Russia  
Tel: 70953306529  
Fax: 70953306538  
Email: sverdshungen.slobc.msk.su.

FEATURES  
Source location/Qualifiers  
1..406  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="ph6h6\_19/1TV"  
/clone\_lib="Outward Alu-primed hncDNA library"  
/note="Vector: pGEM-3Z; Site 1: EcoRI; Site 2: BamHI; The  
library was constructed as described in [Obadovic, D.,  
Borodin, A.M., Kopantzev, E.P., Wagner, L.L., Volik, S.V.,  
Ermolaeva, O.D., Lebedev, Y.B., Monastyrskaya, G.S.,  
Sverdlov, E.D. (1993) Bioorganicheskaya khimiya, 20,  
919-930]. This protocol is based on nested primer strategy  
using Alu- specific primers (ALN3 and TC-65) that direct  
the hncDNA synthesis outward of Alu repeats."

BASE COUNT 80 a 105 c 99 g 122 t

ORIGIN

Alignment Scores:  
Pred. No.: 5.88e-05 Length: 406  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x T48592 (1-406)

QY 875 HisSerAspPheIleValIstYrArgIyValIserTyrGlyProGly 890  
|||||  
DB 25 CACAGTGAATTCATGTCATGATGTCGACCTATGCCCGAGGT 72

RESULT 74  
BM147781 430 bp mRNA linear EST 30-NOV-2001  
LOCUS TCAAP1014273 Pediatric acute myelogenous leukemia cell (FAB M1)  
DEFINITION Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP1427, mRNA  
sequence.  
ACCESSION BM147781.1 GI:17167905  
VERSION  
KEYWORDS EST.

SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R., Jr.,  
Gunnarone,P.H., Munzy,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.  
TITLE Pediatric Leukemia cDNA Sequencing Project (2001)  
JOURNAL Unpublished (2001)  
COMMENT Contact: Dr. Judith F. Margolin  
Texas Children's Cancer Center and Human Genome Sequencing Center  
at Baylor College of Medicine  
1102 Bates, MC3-3320 Houston, TX 77030, USA  
Tel: 832-825-4536  
Fax: 832-825-4038  
Email: clones@tccc.org  
Seq primer: M13 primer.

FEATURES  
Source location/Qualifiers  
1..430  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="TCAAP1427"  
/clone\_lib="pediatric acute myelogenous leukemia cell (FAB  
M1) Baylor-HGSC project-TCAA"  
/sex="male"  
/tissue\_type="leukopheresis"  
/cell\_type="myeloid cell"  
/dev\_stage="pediatric 6 years"  
/lab\_host="DH10B"  
/note="Vector: lambda PSB; Site 1: BamHI; Site 2: EcoRI;  
First strand cDNA was primed with an anchored  
XhoI-oligo(dT) primer [5'GGAGACTGAGCGCGCGAGAGAG(T)VN  
3'; V-A,C,G; N-A,C,G,T] and then dg tailed. Second strand  
was primed with a BamHI-dc primer  
[5'AGAGAGCTGAGTCGCGCGCGCAATATATAT(C) 3'].  
Double-stranded cDNA was then digested with BamHI and XhoI  
and directionally cloned into the BamHI and SalI sites of  
lambda PSB vector. Library went through one round of  
normalization. Library was constructed by Wei Yu at RIKEN  
of Japan (Carinaci P, Westover A, Nishiyama Y, Ohsumi T,  
Itoh M, Nagaoka S, Sasaki, Okazaki Y, Muramatsu M,  
Schneider C, Hayashizaki Y, High efficiency selection of  
full-length cDNA by improved biotinylated cap trapper,  
DNA Res 4: 1, 61-6, Feb 28, 1997)"

BASE COUNT 59 a 149 c 126 g 96 t

ORIGIN

Alignment Scores:  
Pred. No.: 6.29e-05 Length: 430  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 13 Gaps: 0

US-09-397-967-16 (1-1099) x BM147781 (1-430)

QY 62 GlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluasp 77  
|||||  
DB 248 GGCATCTGCGCTGTACACCTCCTCTTGTCTGGCAGCAGAGAC 295

RESULT 75  
BG364706 532 bp mRNA linear EST 08-MAR-2001  
LOCUS dc75a07.y1 NICHD XGC Embl Xenopus laevis cDNA clone IMAGE:3402900  
DEFINITION 5' similar to TR-019064 019064 JMK2. ;, mRNA sequence.  
ACCESSION BG364706  
VERSION BG364706.1 GI:13253803  
KEYWORDS EST.  
SOURCE African clawed frog.  
ORGANISM Xenopus laevis  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;



BASE COUNT 249 a 153 c 188 g 163 t

ORIGIN

Alignment Scores:

Pred. No.:	0.000122	Length:	753
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	12	Gaps:	0

US-09-397-967-16 (1-1099) x BG870416 (1-753)

Oy 1003 SeraspValTrpSerPheGlyValValLeuTyrgluLeuPheThrTyr 1018

DB 397 TCAGATGCTGTGAGCTTTGGAGTGTCTATACGAACTTTTCACATAC 444

RESULT 78

LOCUS B1153342 766 bp mRNA linear EST 05-JUL-2001

DEFINITION 602918614F1 NCI\_CGAP\_Lu29 Mus musculus cDNA IMAGE:5068946 5', mRNA sequence.

ACCESSION B1153342

VERSION B1153342.1 GI:14613343

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE NIH-MGC http://mgc.nci.nih.gov/

AUTHORS 1 (bases 1 to 766)

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLM11185 row: 1 column: 03  
High quality sequence stop: 760.

FEATURES

source

1..766

/organism="Mus musculus"

/strain="CZECH II"

/db\_xref="taxon:10090"

/clone="IMAGE:5068946"

/clone\_1lb="NCI\_CGAP\_Lu29"

/tissue\_type="spontaneous tumor, metastatic to mammary."

/lab\_host="DHI0B"

/stem\_cell\_origin=""

/note="Organ: Lung; Vector: PCMV-SPORT6; Site: 1: SalI; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 254 a 145 c 194 g 173 t

ORIGIN

Alignment Scores:

Pred. No.:	0.000125	Length:	766
Score:	16.00	Matches:	16

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.46%

DB: 13

US-09-397-967-16 (1-1099) x B1153342 (1-766)

Oy 1003 SeraspValTrpSerPheGlyValValLeuTyrgluLeuPheThrTyr 1018

DB 444 TCAGATGCTGTGAGCTTTGGAGTGTCTATACGAACTTTTCACATAC 491

RESULT 79

LOCUS B0899824 920 bp mRNA linear EST 16-AUG-2002

DEFINITION AGENCOURT\_8755626 NIH\_MGC\_130 Mus musculus cDNA clone IMAGE:6334890 5', mRNA sequence.

ACCESSION B0899824

VERSION B0899824.1 GI:22291838

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE NIH-MGC http://mgc.nci.nih.gov/

AUTHORS 1 (bases 1 to 920)

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Mark Macconochie, Ph.D. and Nancy L. Freeman, Ph.D.  
CDNA Library Preparation: ResGen, Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLM13796 row: 9 column: 19  
High quality sequence stop: 663.

FEATURES

source

1..920

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:6334890"

/clone\_1lb="NIH\_MGC\_130"

/lab\_host="DHI0B (phage-resistant)"

/note="Organ: otcocysts; Vector: PCMV-SPORT6.1.ccd; Site: 1: EcoRV; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH-MGC Library."

BASE COUNT 295 a 192 c 216 g 215 t 2 others

ORIGIN

Alignment Scores:

Pred. No.:	0.000155	Length:	920
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	14	Gaps:	0

US-09-397-967-16 (1-1099) x B0899824 (1-920)

Oy 1003 SeraspValTrpSerPheGlyValValLeuTyrgluLeuPheThrTyr 1018

DB 288 TCAGATGCTGTGAGCTTTGGAGTGTCTATACGAACTTTTCACATAC 335

RESULT 80

LOCUS B0732700 945 bp mRNA linear EST 16-JUL-2002

DEFINITION AGENCOURT\_8097616 NICHD XGC Emb4 Xenopus laevis cDNA clone IMAGE:5572666 5', mRNA sequence.

ACCESSION B0732700.1 GI:21871597  
 VERSION EST.  
 KEYWORDS African clawed frog.  
 SOURCE XENOPUS laevis  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.  
 REFERENCE 1 (bases 1 to 945)  
 AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), National Cancer Institute  
 JOURNAL Tumor Gene Index  
 COMMENT Unpublished (1997)  
 CONTACT Robert Strausberg, Ph.D.  
 EMAIL [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 TISSUE Procurement: Dr. Igor David  
 CDNA LIBRARY PREPARATION: Life Technologies, Inc.  
 CDNA LIBRARY ARRAYED BY: The I.M.A.G.E. Consortium (LLNL)  
 DNA SEQUENCING BY: Agencourt Bioscience Corporation  
 CLONE DISTRIBUTION: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www.bio.llnl.gov/bdrp/image/image.html](http://www.bio.llnl.gov/bdrp/image/image.html)  
 PLATE: L1AM12318 row: h column: 11  
 HIGH QUALITY SEQUENCE STOP: 352.  
 FEATURES  
 SOURCE  
 1..945  
 Location/Qualifiers  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 /clone="IMAGE:5572666"  
 /clone\_1lb="NICHD XGC Embd4"  
 /dev\_stage="embryo, stage 31-32"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: whole embryo; Vector: pCMV-SPORT6; Site: 1; Note: Site 2: Salt; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."  
 BASE COUNT 315 a 197 c 217 g 215 t 1 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.00016 Length: 945  
 Score: 16.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: 14 Gaps: 0  
 US-09-397-967-16 (1-1099) x B0732700 (1-945)  
 QY 1003 SerAspValTrpSerPheGlyValValLeuTyrgluLeuPheThrTyr 1018  
 |||||||  
 Db 304 TCGAGTCGTGGAGCTTGGAGTTGTTCTCTATGAGTGTTCACATAC 351  
 RESULT 81  
 CENS042XQ 989 bp DNA linear GSS 18-MAY-2000  
 LOCUS Tetraodon nigroviridis genome survey sequence PUC-ori end of clone  
 DEFINITION 078611 of library G from Tetraodon nigroviridis, genomic survey  
 sequence.  
 ACCESSION AL272087.1 GI:7994146  
 VERSION GSS: genome survey sequence.  
 KEYWORDS Tetraodon nigroviridis.  
 SOURCE Tetraodon nigroviridis.  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.  
 REFERENCE 1 (bases 1 to 989)  
 AUTHORS Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.

TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 989)  
 AUTHORS Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.  
 TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 989)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-APR-2000)  
 COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.  
 FEATURES  
 SOURCE  
 1..989  
 Location/Qualifiers  
 /organism="Tetraodon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="078f11"  
 /clone\_1lb="G"  
 /note="Genoscope sequence ID : C0BG078C06SP1-end : PUC-ori"  
 BASE COUNT 237 a 288 c 275 g 184 t 5 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.000169 Length: 989  
 Score: 16.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: 17 Gaps: 0  
 US-09-397-967-16 (1-1099) x CENS042XQ (1-989)  
 QY 1003 SerAspValTrpSerPheGlyValValLeuTyrgluLeuPheThrTyr 1018  
 |||||||  
 Db 414 TCCGACCTTGGAGCTTGGCGCTGCTCTACGAGCTTTCACCTAC 461  
 RESULT 82  
 CENS038MW/C 1001 bp DNA linear GSS 15-MAY-2000  
 LOCUS Tetraodon nigroviridis genome survey sequence PUC-ori end of clone  
 DEFINITION 005423 of library G from Tetraodon nigroviridis, genomic survey  
 sequence.  
 ACCESSION AL232817.1 GI:7891952  
 VERSION GSS: genome survey sequence.  
 KEYWORDS Tetraodon nigroviridis.  
 SOURCE Tetraodon nigroviridis.  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.  
 REFERENCE 1 (bases 1 to 1001)  
 AUTHORS Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.  
 TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1001)  
 AUTHORS Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.  
 TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 1001)

AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (12-APR-2000)  
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.

FEATURES  
source  
1. 1001  
/organism="Tetraodon nigroviridis"  
/db\_xref="taxon:99883"  
/clone\_lib="g"  
/note="Genoscope sequence ID : COBG005CE12XE1-end ;  
PUC-ori"

BASE COUNT 209 a 265 c 274 g 249 t 4 others

ORIGIN

Alignment Scores:  
Pred. No.: 0.000172 Length: 1001  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 17 Gaps: 0

US-09-397-967-16 (1-1099) x CNS038MM (1-1001)

OY 1003 Serpaspaltrpserpghlyvalvalleuryglureupherhrrtyr 1018  
|||||  
Db 273 TCCGACGTTTGAGCTCGGCGTCTCTACGAGCTCTACCTAC 226

RESULT 83  
BF373713/ 253 bp mRNA linear EST 24-NOV-2000  
LOCUS BF373713 PM4-SN0016-080600-005-b03 SN0016 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BF373713  
ACCESSION BF373713  
VERSION BF373713.1 GI:11335829  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 253)  
AUTHORS Nagai, N., Garcia, R., Verjovski-Almeida, S., Briones, M.R., Naga, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl-PM4&lt2-PM4-SN0016-080600-005-b03&lt3-2000-06-08&lt4-1>)  
Seq primer: puc 18 forward  
High quality sequence start: 17  
High quality sequence stop: 66.  
Location/Qualifiers  
1. 253  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

FEATURES  
source

/clone\_lib="SN0016"  
/dev\_stage="Adult"  
/note="Organ: stomach, normal; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 58 a 79 c 60 g 56 t

ORIGIN

Alignment Scores:  
Pred. No.: 0.000319 Length: 253  
Score: 15.00 Matches: 15  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.36% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967-16 (1-1099) x BF373713 (1-253)

OY 600 Leuglyalalleaspmetrytleuarglyargglyhisleuval 614  
|||||  
Db 240 CTGGGGCCATAGACATGTATCTGGAAACGTGCCACCTGTTG 196

RESULT 84  
BH043390 754 bp DNA linear GSS 17-JUN-2001  
LOCUS BH043390  
DEFINITION RPCI-24-360L21.TJ RPCI-24 Mus musculus genomic clone RPCI-24-360L21  
ACCESSION BH043390  
VERSION BH043390.1 GI:14825476  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 754)  
AUTHORS Zhao, S., Nieman, W., Malek, J., Shatsman, S., Akincet, B., Levins, M., Tesgaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregeorgis, E., Russell, D., de Jong, P. and Fraser, C.M.  
TITLE Mouse BAC End Sequences from Library RPCI-24  
JOURNAL Unpublished (1999)  
COMMENT Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: [szhao@tigr.org](mailto:szhao@tigr.org)  
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong ([pedejong@tigr.org](mailto:pedejong@tigr.org)). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: [http://www.tigr.org/cdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 360 row: L column: 21  
Seq primer: SP6  
Class: BAC ends.  
Location/Qualifiers  
1. 754  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-24-360L21"  
/clone\_lib="RPCI-24"  
/sex="Male"  
/cell\_type="Spleen/Brain"  
/note="Vector: pYRABAC1; Site\_1: BamHI; Site\_2: BamHI; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pYRABAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

FEATURES  
source



BASE COUNT 144 a 224 c 210 g 176 t  
 ALIGNMENT SCORES:  
 Pred. No.: 0.00117 Length: 754  
 Score: 15.00 Matches: 15  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.36% Indels: 0  
 DB: 17 Gaps: 0  
 US-09-397-967-16 (1-1099) x BH043390 (1-754)  
 QY 333 GUAAlauserPheValAlaLeuValAspGlyTyrPheArgLeu 347  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 584 GAGGCCCTGCTTGTGCGCCGTGCGATGCTATTTCGCGTGTG 628  
 RESULT 85  
 LOCUS AM437524 281 bp mRNA linear EST 25-APR-2001  
 DEFINITION 78773 MARC IBOV Bos taurus cDNA 5', mRNA sequence.  
 ACCESSION AM437524  
 VERSION AM437524.1 GI:6972830  
 KEYWORDS EST.  
 SOURCE cow.  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 1 (bases 1 to 281)  
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
 Caass,E., Wray,J.E., White,J., Cho,J., Fahnenkrug,S.C., Bennett,  
 G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,  
 Perlea,G., Holt,I., Karameycheva,S., Liang,F., Quackenbush,J. and  
 Keeler,J.W.  
 Sequence evaluation of four pooled-tissue normalized bovine cDNA  
 libraries and construction of a gene index for cattle  
 Genome Res. 11 (4), 626-630 (2001)  
 JOURNAL  
 MEDLINE  
 COMMENT  
 CONTACT: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smitht@mail.marc.usda.gov  
 Single pass sequencing. Bases called and trimmed with phred  
 v0.980904.e. Vector identified by cross-match with the -minscore 20  
 and -minmatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCATGACCAT  
 BACKWARD: GTTTCACGTCACGACG  
 Plate: 42 row: A column: 9  
 Seq primer: ATTTAGGTGACATATAG.  
 FEATURES  
 SOURCE  
 Location/Qualifiers  
 1..281  
 /organism="Bos taurus"  
 /db\_xref="taxon:9913"  
 /clone\_lib="MARC IBOV"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV SPORT6; site\_1: NotI; site\_2: SalI;  
 library made from pooled tissue from lymph node, ovary,  
 fat, hypothalamus, and pituitary."  
 BASE COUNT 41 a 95 c 91 g 54 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.00345 Length: 281  
 Score: 14.00 Matches: 14  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.27% Indels: 0  
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x AM437524 (1-281)  
 QY 334 AlaLeuserPheValAlaLeuValAspGlyTyrPheArgLeu 347  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 58 GCGTTGTCTTGTGCGCGCTGCGATGCGTACTTCCGCGTGTG 99  
 RESULT 86  
 LOCUS AM212845/c 296 bp mRNA linear EST 03-DEC-1999  
 DEFINITION u068d02.x1 NCI CGAP Mam1 Mus musculus cDNA clone IMAGE:2647683.3,  
 similar to gb:U40172 Mus musculus JAK3 gene, complete cds (MOUSE);,  
 mRNA sequence.  
 ACCESSION AM212845  
 VERSION AM212845.1 GI:6518985  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 296)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Other ESTs: u068d02.y1  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapdb-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
[www-bio.lindl.gov/bdrp/image/image.html](http://www-bio.lindl.gov/bdrp/image/image.html)  
 MGI:1028135  
 Seq primer: -40UP from gibco.  
 FEATURES  
 SOURCE  
 Location/Qualifiers  
 1..296  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_image="IMAGE:2647683"  
 /clone\_lib="NCI CGAP Mam1"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="10 months, virgin"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: PCMV-SPORT6; site\_1: SalI;  
 site\_2: NotI; Cloned unidirectionally. Primer: Oligo dr.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"  
 BASE COUNT 62 a 71 c 95 g 68 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.00367 Length: 296  
 Score: 14.00 Matches: 14  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.27% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-397-967-16 (1-1099) x AM212845 (1-296)  
 QY 1062 ProThrGluValGlnGluLeuMetGlnLeuGlyTrpAlaPro 1075  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 295 CCCACCGAGGTTCAGAGACTCATGCACTGTGTGCGCCGCC 254  
 RESULT 87  
 LOCUS AA008555 318 bp mRNA linear EST 25-JUL-1996  
 DEFINITION mg86b10.r1 Soares mouse embryo NMME13.5 14.5 Mus musculus cDNA

containing the following sequence: 5'-GGAGGAGACATTCACAGACCTTTTGTGTTTNN 3'}. cDNA was primed with a primer [5'] GGAGGAGACATTCACAGACCTTTTGTGTTTNN 3'} prepared by using trehalase thermo-activated reverse transcriptase and subsequently enriched for full-length transcripts by PCR amplification with the cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'] GAGGAGACATTCGAGTAAATTGAATTATCATCCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified

BASE COUNT 156 a 85 c 104 g 104 t  
ORIGIN bluescript KS(+) after bulk excision from Lambda FLC I."

## Alignment Scores:

Pred. No.: 0.00601 Length: 449  
Score: 14.00 Matches: 14  
Percent Similarity: 100.00% Conservatave: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.27% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x BB711016 (1-449)

OY 1005 ValTSPePheGlyValValLeuTyrGluLeuPheThrTyr 1018  
DB 359 CTTGTGAGCTTGTGAGCTGTCTATACGACTTTTCACATAC 400

## RESULT 89

BO780982/c

LOCUS 636 bp mRNA linear EST 26-JUL-2002  
DEFINITION UI-R-FFO-cpd-c-08-0-UI.s1 UI-R-FFO Rattus norvegicus cDNA clone

ACCESSION BO780982  
VERSION BO780982.1 GI:21989454

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 636)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT Contact: Soares, MB

Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250  
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

Tissue Procurement: Jeff Stevens  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: DISTRIBUTION: Researchers may obtain clones

from Research Genetics (www.resgen.com).  
Seq primer: M13 FORWARD

POLYA=Yes.  
Location/Qualifiers

FEATURES  
source

1. 636  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="UI-R-FFO-cpd-c-08-0-UI"  
/clone\_lib="UI-R-FFO"  
/tissue\_type="Mixed tissues"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Vector: p773-Pac (Pharmacia) with a modified  
polylinker. Site\_1: EcoR I; Site\_2: Not I; UI-R-FFO is a  
subtracted cDNA library containing the following tissue(s)  
): Normal cartilage and SR-JMS Tumor Line. The  
subtraction was made according to Bonaldo, Lennon and  
Soares, Genome Research, 6:791-806, 1996. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for these libraries are: CTAATGACG,  
CATCTGTGA,  
TAG\_LIB-UI-R-FFO

TAG\_TISSUE=cartilage  
TAG\_SEQ=CTAATGACG"  
BASE COUNT 139 a 156 c 200 g 138 t 3 others  
ORIGIN

## Alignment Scores:

Pred. No.: 0.00909 Length: 636  
Score: 14.00 Matches: 14  
Percent Similarity: 100.00% Conservatave: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.27% Indels: 0  
DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x BO780982 (1-636)

OY 1045 LeuLeuGluLeuLeuAgluGlyArgLeuProPro 1058  
DB 609 CTTGTGAGCTGTGAGGAGGCGCCCTCTCCACACCC 568

## RESULT 90

BO781477/c

LOCUS 683 bp mRNA linear EST 26-JUL-2002  
DEFINITION UI-R-FFO-cpg-e-19-0-UI.s1 UI-R-FFO Rattus norvegicus cDNA clone

ACCESSION BO781477  
VERSION BO781477.1 GI:21989494

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 683)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT Contact: Soares, MB

Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250  
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

Tissue Procurement: Jeff Stevens  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: DISTRIBUTION: Researchers may obtain clones

from Research Genetics (www.resgen.com).  
Seq primer: M13 FORWARD

POLYA=Yes.  
Location/Qualifiers

FEATURES  
source

1. 683  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="UI-R-FFO-cpg-e-19-0-UI"  
/clone\_lib="UI-R-FFO"  
/tissue\_type="Mixed tissues"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Vector: p773-Pac (Pharmacia) with a modified  
polylinker. Site\_1: EcoR I; Site\_2: Not I; UI-R-FFO is a  
subtracted cDNA library containing the following tissue(s)  
): Normal cartilage and SR-JMS Tumor Line. The  
subtraction was made according to Bonaldo, Lennon and  
Soares, Genome Research, 6:791-806, 1996. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for these libraries are: CTAATGACG,  
CATCTGTGA.



vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. Plants were inoculated by Shuxian Li (Glen Hartman Lab, University of Illinois). Library was constructed by Reena Phillip and Steve Clough (Lila Vodkin Lab, University of Illinois)."

BASE COUNT 156 a 90 c 64 g 131 t  
ORIGIN

## Alignment Scores:

Pred. No.:	0.0561	Length:	441
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.18%	Indels:	0
DB:	12	Gaps:	0

US-09-397-967-16 (1-1099) x BG509232 (1-441)

Oy 1003 SeraspvairpserphgilyvalleuTYrGluLeu 1015

DB 133 AGTGATGTGTGAGCTTGTGTAGTCTTATGAGCTT 95

## RESULT 93

BM813523 444 bp mRNA linear EST 05-MAR-2002  
LOCUS BM813523  
DEFINITION EST591616 BNR Medicago truncatula/Meloidogyne incognita mixed EST  
library cDNA clone pBNIR-10G16, mRNA sequence.

ACCESSION BM813523.1 GI:19149537

## VERSION

BM813523.1 Medicago truncatula/Meloidogyne incognita mixed EST library.

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

1. 444  
Location/Qualifiers  
/organism="Medicago truncatula/Meloidogyne incognita mixed EST library"  
/cultivar="A17"  
/db\_xref="taxon:188702"  
/clone="pBNIR-10G16"  
/clone.lib="BNIR"  
/tissue\_type="3 week old roots"  
/dev\_stage="3 days after infection with Meloidogyne incognita second stage larvae"  
/lab\_host="XLOLR"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Strategene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."

## BASE COUNT

## ORIGIN

## Alignment Scores:

Pred. No.: 0.0565 Length: 444

Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x BM813523 (1-444)

Oy 1003 SeraspvairpserphgilyvalleuTYrGluLeu 1015

DB 24 AGTGATGTGTGAGCTTGTGTAGTCTTATGAGCTT 62

## RESULT 94

BI424599 500 bp mRNA linear EST 29-NOV-2001  
LOCUS BI424599  
DEFINITION saH53602.Y1 Gm-c1036 Glycine max cDNA clone GENOME SYSTEMS CLONE  
ID: Gm-c1036-5188-5' similar to TR:081064 081064 HYPOTHETICAL 51.9  
KDA PROTEIN. ; mRNA sequence.

ACCESSION BI424599.1 GI:15200920

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

1 (bases 1 to 500)  
Shoemaker, R., Kelm, P., Vodkin, L., Eprelding, J., Corry, V., Khanna, A., Bolla, B., Maria, M., Hillier, L., Kueba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Willson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: cuteresgen.com  
High quality sequence stop: 414.

## FEATURES

## source

1. 500  
Location/Qualifiers  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1036-5188"  
/clone.lib="Gm-c1036"  
/tissue\_type="somatic embryos cultured on MSD 20"  
/lab\_host="DH10B"  
/note="Vector: pSPORT1; Site\_1: NotI; Site\_2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the life technologies superscript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linker adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

## BASE COUNT

## ORIGIN

130 a 90 c 136 g 144 t

Alignment Scores:

Pred. No.:	0.0651	Length:	500
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.18%	Indels:	0
DB:	13	Gaps:	0

US-09-397-967-16 (1-1099) x B1424599 (1-500)

QY 1003 SeraspValTrpSerPheGlyValValLeuTyrgIleu 1015  
 |||||||  
 DB 348 AGTGATGTGTGAGCTTGTGTAGTCTTATGAGCTT 386

RESULT 95  
 AM231037 562 bp mRNA linear EST 10-DEC-1999  
 LOCUS uc07a12.y1 NCI-CGAP\_Mam1 Mus musculus cDNA IMAGE:2647870 5'  
 DEFINITION similar to SW:JAK2\_MOUSE Q62120 TYROSINE-PROTEIN KINASE JAK2 ;  
 mRNA sequence.  
 ACCESSION AM231037.1 GI:6560269  
 VERSION EST.  
 KEYWORDS house mouse;  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 562)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 www.bio.lnl.gov/bdrr/image/image.html

MG1:1028322  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 402.

FEATURES  
 source  
 1..562  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone\_image="IMAGE:2647870"  
 /clone\_1lb="NCI-CGAP\_Mam1"  
 /tissue\_type="tumor biopsy sample"  
 /dev\_stage="10 months, virgin"  
 /lab\_host="DRI0B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dr.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"

BASE COUNT 181 a 114 c 147 g 120 t

ORIGIN

Alignment Scores:

Pred. No.:	0.0748	Length:	562
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.18%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-16 (1-1099) x AM231037 (1-562)

QY 1003 SeraspValTrpSerPheGlyValValLeuTyrgIleu 1015

DB 524 TCAGATGTGTGAGCTTGTGAGTGTCTCATACGAGCTT 562  
 |||||||  
 RESULT 96  
 BE156215/c  
 LOCUS BE156215  
 DEFINITION OVO-HT0367-310100-102-h08 HT0367 Homo sapiens cDNA, mRNA sequence.  
 BE156215  
 ACCESSION BE156215.1 GI:6618936  
 VERSION EST.  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 145)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brites,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-QVO-HT0367-310  
 100-102-H08&t3=2000-01-31&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 17  
 High quality sequence stop: 145.

FEATURES  
 source  
 1..145  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_1lb="HT0367"  
 /dev\_stage="Adult"  
 /note="Organ: head,neck; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORSTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 35 a 40 c 42 g 28 t

ORIGIN

Alignment Scores:

Pred. No.:	0.143	Length:	145
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-16 (1-1099) x BE156215 (1-145)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 |||||||  
 DB 91 GTGCACCGCGACCTGGCTGCTGCACACATCCTTCTC 56

RESULT 97  
 AA320481  
 LOCUS AA320481 214 bp mRNA linear EST 19-APR-1997

```

IMAGE:160664.5' similar to gb:M57464 PROTO-ONCOGENE
TYROSINE-PROTEIN KINASE RECEPTOR RET PRECURSOR (HUMAN) ; mRNA
sequence.
H24996 GI:893895
human.
EST.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 270)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,T.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 809
High quality sequence stops: 166
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAG Consortium (info@image.lnl.gov) for further information.
Insert Length: 809 Std Error: 0.00
Seq primer: M13Rp1
High quality sequence stop: 166.
FEATURES
source
Location/Qualifiers
1..270
/organism="Homo sapiens"
/db_xref="GB:574726"
/db_xref="taxon:9606"
/clone="IMAGE:160664"
/clone_lib="Soares breast 3MBdbst"
/dev_stage="Female"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: breast; Vector: pT73D (Pharmacia) with a
modified polylinker. Site_1: Not I; site_2: Eco RI; 1st
strand cDNA was primed with a Not I - Oligo(df) primer [5'
TGTTACCAATCTGAACTGGAGCGGCCCGCCCTTTTTTTTTTTTTTTT 3'] ,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of a modified pT73 vector (Pharmacia).
Library went through one round of normalization to a Cot =
20. Library constructed by Bento Soares and M.Patima
Bonaldo."
BASE COUNT 65 a 52 c 74 g 69 t 10 others
ORIGIN
Alignment Scores:
Pred. No.: 0.298 Length: 270
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatch: 0
Query Match: 1.09% Indels: 0
Gaps: 0
US-09-397-967-16 (1-1099) x H24996 (1-270)
QY 941 ValHisArgAspLeuAlaAlaIArgAsnIleLeuVal 952
Db 29 GTTCATCGGACTGTGCAGCAGAACAATCCTGGTA 64
LOCUS B1051048 294 bp mRNA linear EST 15-JUN-2001
DEFINITION CM2-GN0283-100101-683-g12 GN0283 Homo sapiens cDNA, mRNA sequence.
ACCESSION B1051048

```

```

VERSION      B01051048.1  GI:14458578
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE    1 (bases 1 to 294)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
              Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
              Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
              Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
              M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
              Simpson,A.J.
              Shotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
              Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL      20202663
MEDLINE
COMMENT      Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research
              Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
              Brazil
              Tel.: +55-11-2704922
              Fax: +55-11-2707001
              Email: asimpson@ludwig.org.br
              This sequence was derived from the FAPESP/LICR Human Cancer Genome
              Project. This entry can be seen in the following URL
              (http://www.ludwig.org.br/scripts/gethtml2.pl?l=CM2&l2=CM2-GN0283-100101-683-gj2&t3=2001-01-10&t4=1)
              Seq primer: puc 18 forward
              High quality sequence start: 14
              High quality sequence stop: 294.
              Location/Qualifiers
                1..294
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone_id="GN0283"
                  /dev_stage="Adult"
                  /note="Organ: Placenta, normal; Vector: puc18; Site_1: SmaI
                  ; Site_2: SmaI; A mini-library was made by cloning
                  products derived from ORESTES PCR (U.S. Letters Patent
                  application No. 196,716 - Ludwig Institute for Cancer
                  Research) profiles into the pUC 18 vector. Reverse
                  transcription of tissue mRNA and cDNA amplification were
                  performed under low stringency conditions."
BASE COUNT   58 a      81 c      90 g      65 t
ORIGIN
Alignment Scores:
Pred. No.:      0.33      Length:      294
Score:          12.00      Matches:      12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    1.09%      Indels:      0
DB:             13      Gaps:      0
US-09-397-967-16 (1-1099) x B1051048 (1-294)
Oy      941  ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
          |||||||
          270  GTTCACCGTCGACCTGGCTGCCGACACATCTCTGTC 235
          |||||||
RESULT 100
BB605196      308 bp      mRNA      linear      EST 05-DEC-2000
LOCUS        BB605196      RIKEN full-length enriched, 0 day neonate lung Mus
DEFINITION   musculus cDNA clone E030009B22 5', mRNA sequence.
ACCESSION    BB605196
VERSION      BB605196.1  GI:11556598
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 308)	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus			
Alizawa, K., Akahira, S., Akimura, T., Arai, A., Arakawa, T., Carninci, P.,				
Hayatsu, N., Hirotsu, T., Hirozane, T., Hodojima, Y.,				
Imochani, K., Ishii, Y., Itoh, M., Izawa, M., Kawai, J., Kodama, Y.,				
Kusakabe, M., Matsuyama, T., Miyazaki, A., Nakamura, M., Nishi, K.,				
Nomura, K., Nunazaki, R., Okazaki, Y., Okido, T., Owa, C., Sakai, C.,				
Sakai, K., Sasaki, D., Sato, K., Shibata, K., Shibata, Y., Shinagawa, A.,				
Shiraki, T., Sugabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Tanaka				
T., Toyota, T., Wataniki, A., Yamamura, T., Yasunishi, A., Yoshida, K.,				
Yoshiki, A., Yamatsu, M. and Hayashizaki, Y.				
RIKEN Mouse ESTs (Alizawa, K. et al. 2000)				
Unpublished (2000)				
contact: Yoshhide Hayashizaki				
laboratory for Genome Exploration Research Group, RIKEN genomic				

COMMENT

Contact: Toshinori Hayashi  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.go.jp,  
url: <http://genome.gsc.riken.go.jp/>  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasak  
, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermotabilization and thermocactivation of the thermostable enzymes b  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (22) 520-524 (1998)  
Itoh, M., Kitsuina, T., Akiyama, J., Shibata, K., Irawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, T., Muramatsu, M., Okazak  
, Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
further details.

```

FEATURES
Source
Location/Qualifiers
1. .308
/organism="Mus musculus"
/db_xref="taxon:10090"
/cclone="E030009B22"
/clone_1lb="RIKEN full-length enriched, 0 day neonate lung"
/tissue-type="lung"
/dev-stage="0 day neonate"
/lab_host="DH10B"
/note="Site-1: SalI; Site-2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15' GAGGAGAGACGGCCGCCAATCTCGAGTGTTCCTTTTTCCTTTCVN 3'}, cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length transcripts. Second strand cDNA was prepared with the cap-trapper. Primer adapter of sequence 15' GAGGAGAGATTCTCGATTATTAATTAATTAATCCCCCCCCCCC 3'} cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLX I."
BASE COUNT
76 a 79 c 93 g 60 t
ORIGIN
Alignment Scores:
Pred. No.: 0.349 Length: 308
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatch: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

```



us-09-397-967-16.oligo.rst

Page 53

QY	941	ValHisArgAspLeuAlaAlaArgAsnIleLeuVal	952
Db	122	GTGCACCCGGACCTGGCCGCCGCAACATCTCTGTC	157

Search completed: April 28, 2003, 21:38:48  
Job time : 3067 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## SUMMARIES

OM protein - nucleic acid search, using frame\_plus\_p2n model

Run on: April 28, 2003, 16:55:25 ; Search time 524 Seconds  
(without alignments)  
4723.180 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 1099

Sequence: 1 MAPSEEMPPLIPQRSCSLSS.....RPAFAVILSPOLDPLWRGRPG 1099

Scoring table:

OLIGO  
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4368727

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 150 summaries

Command line parameters:

MODEL=frame\_plus\_p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO-spool/US09397967/runat\_26042003\_182926\_11190/app\_query.fasta.1.1287  
-DB=N.Geneseq.101002 -OFMT=fastp -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -FRANS=human40.cdd  
-LIST=150 -DOCALLIGN=200 -THR SCORE=quality -THR\_MIN=1 -ALIGN=100 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09397967.ecgn.1.1.362-etunat\_26042003\_182926\_11190 -NCPU=6 -ICPU=3  
-NO\_XLPRX -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: N.Geneseq.101002.\*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	180	16.4	4016	17	AAT11083	Mouse JAK3 protein
2	62	5.6	3807	17	AAT30862	Protein tyrosine k
3	51	4.6	795	22	AA510808	Human Janus kinase
4	40	3.6	778	22	AA510809	Human Janus kinase
5	31	2.8	165	16	AA087512	Tyrosine kinase pe
6	27	2.6	811	22	AA510806	Human Janus kinase
7	29	2.5	347	24	ABN21423	Human ORFX polynuc
8	23	2.1	3222	22	AAK78095	Human ORFX polynuc
9	21	1.9	65	24	ABN5677	Human Immune/haema
10	18	1.6	109	22	AA510804	Mouse spliced tran
11	18	1.6	726	22	AA510807	Human Janus kinase
12	18	1.6	2000	22	AAK78097	Human Immune/haema
13	16	1.5	3435	19	AAV61801	Human Janus kinase
14	16	1.5	3473	13	AA025307	JAK2 protein encod
15	16	1.5	3495	24	AAAD2680	JAK2 encoding DNA.
16	16	1.5	3629	16	AA085412	Murine JAK2 tyrosi
17	16	1.5	3629	21	AAAC6244	Murine JAK2 kinase
18	16	1.5	3629	22	AAAD3607	JAK2 polynucleotid
19	16	1.5	4482	20	AAK80971	Murine (Janus) kina
20	16	1.5	4482	21	AAZ58947	Human JAK2 kinase
21	16	1.5	5117	24	AAAD2431	Human JAK2 kinase
22	14	1.3	1498	22	AAK78096	Human JAK2 (Janus
23	12	1.1	151	14	AAQ49750	Human Immune/haema
24	12	1.1	151	16	AAAT03093	PTK gene LPTK3 par
25	12	1.1	214	22	AA510805	Protein tyrosine-k
26	12	1.1	294	22	ABA50838	Human Janus kinase
27	12	1.1	294	22	ABA68808	Human breast cell
28	12	1.1	294	22	ABA35766	Human foetal liver
29	12	1.1	294	22	AAK17149	Probe #14232 for g
30	12	1.1	294	22	AAK42934	Human brain expres
31	12	1.1	294	22	AAI23695	Human bone marrow
32	12	1.1	294	22	AAI49009	Probe #13628 for g
33	12	1.1	294	22	AAI09314	Probe #1695 used t
34	12	1.1	294	22	AB516994	Probe #9305 used t
35	12	1.1	367	22	ABA45718	Human genome-deriv
36	12	1.1	367	22	ABA56224	Human breast cell
37	12	1.1	367	22	ABA52869	Human foetal liver
38	12	1.1	367	22	AAK04412	Probe #4335 for ge
39	12	1.1	367	22	AAK29908	Human brain expres
40	12	1.1	367	22	AAI14495	Human bone marrow
41	12	1.1	367	22	AAI15871	Probe #4428 for ge
42	12	1.1	367	22	AAI04332	Probe #4557 used t
43	12	1.1	367	22	AB504474	Human genome-deriv
44	12	1.1	405	21	AAAS7858	Human RET proto-on
45	12	1.1	917	21	AAAC95082	Cat flea hindgut a
46	12	1.1	1117	23	AA587117	DNA encoding novel
47	12	1.1	1509	14	AA053470	PTK gene B1. Ratt
48	12	1.1	1942	16	AA084888	DNA encoding cytop
49	12	1.1	1987	19	AAV44497	Human mark cDNA.
50	12	1.1	2000	16	AAV00616	Megakaryocyte kina
51	12	1.1	2170	20	AAZ41307	Human normal ovary
52	12	1.1	2489	23	AA585447	DNA encoding novel
53	12	1.1	2784	24	ABL91661	Human polynucleoti
54	12	1.1	2820	16	AA090654	Human polynucleoti
55	12	1.1	2955	24	ABJ91658	Eph-related PTK Ce
56	12	1.1	2955	24	ABJ91678	Human polynucleoti
57	12	1.1	2955	24	AAV02946	Human polynucleoti
58	12	1.1	2962	16	AAV02946	Eph-1-like receptor
59	12	1.1	2964	16	AAV02946	Human polynucleoti
60	12	1.1	2997	24	ABJ91662	Protein p140 cDNA
61	12	1.1	2997	24	ABJ91662	Human polynucleoti
62	12	1.1	3042	24	ABJ91659	Human polynucleoti
63	12	1.1	3105	16	AAAT07308	Human polynucleoti
64	12	1.1	3105	18	AAAT07308	Receptor tyrosine
65	12	1.1	3105	22	AA503812	Mouse Nuk tyrosine
66	12	1.1	3107	23	AA569876	Murine neural kina
67	12	1.1	3107	24	ABJ65818	DNA encoding novel
						Lung cancer relate



Sequence 4016 BP; 742 A; 1203 C; 1235 G; 836 T; 0 other:

## Alignment Scores:

Pred. No.:	1.79e-160	Length:	4016
Score:	180.00	Matches:	529
Percent Similarity:	97.24%	Conservative:	0
Best Local Similarity:	97.24%	Mismatches:	9
Query Match:	16.38%	Indels:	15
DB:	17	Gaps:	0

US-09-397-967-16 (1-1099) x AAT11083 (1-4016)

QY 535 ArgArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuValMetAsp 554  
 DB 2257 AGCGCGAGGTCGTGATGATGATGACACATGACTCGGAAGTCTCTGGAAGGCATGAGAC 2316  
 QY 555 SerArgHisArgAsnCysMetGluSerPheLeuGluAlaSerLeuMetSerGlnVal 574  
 DB 2317 TCCAGACATCGGAACTGAGTCTCTTCTGGAAGCCGAAAGCTTGATGAGCCCAAGTA 2376  
 QY 575 SerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetVal 594  
 DB 2377 TCCTACCCGACCTGCTGTATTACGACAGCGGCTGTCATGGCTGGAGACAGCATCATGCTG 2436  
 QY 595 GlnGluPheValTyrLeuGlyAlaAlaLeuAspMetTyrLeuArgGlySerGlyHisLeuVal 614  
 DB 2437 CAGGAATTTGTGTATCTAGAGCAATTTGACATGTACTCTGGCGCAAGCGTGGCCACTGCTG 2496  
 QY 615 SerAlaSerTyrPheLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGln 634  
 DB 2497 TCAGCCAGCTGGAACTGCAAGGTGACCAAGACACTGGCATATGCCCTTAACTACTTGGAG 2556  
 QY 635 AspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGlyGly 654  
 DB 2557 GACAAAGGCTTCTCTCAGCGCAACGCTCAGCAGCGAAGTGTCTCTGCTGCTGAGGGG 2616  
 QY 655 GlyAspGlyAsnProProthetIleLysLeuSerAspProGlyValSerProthIleVal 674  
 DB 2617 GGTGATGGGAATTCACCTTTTCATTAAAGTGTAGTGTCTGTGTCTGCTCCACTGTGCTG 2676  
 QY 675 SerLeuGlnMetLeuThrAspArgIleProTyrValAlaProGlyLysLeuGlnGluAla 694  
 DB 2677 AGCTGTGAAATGCTCACCCGACAGAAATACCTGGGTGGCCCGCAATGTCTCCAGAGAGCT 2736  
 QY 695 GlnThrLeuCysLeuGluAlaAspLysTyrGlyPheGlyAlaThrThrTyrGluValPhe 714  
 DB 2737 CAGACACTGCTGCTTGAAGCTGACAGTGGGCTTGGAGCCACACGCTGGAGGCTGT-- 2795  
 QY 715 GlnArg-GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGly 734  
 DB 2796 CACCGGGGAGACCCGACCATTCACCTCGCTGGAGCCCGCAAAAGCTGAAGTTCTATGA 2855  
 QY 734 uAspGlnGlyLeuProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCys 754  
 DB 2856 GGAACCGAGACACTCCCGCTCTCAATGAGACAGAACTGGCGGACTTATATCACAGAGT 2915  
 QY 754 sMetAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGly 774  
 DB 2916 CAGGGGTATGATCTGGCGGGCGCCCTCTCCGAGATATCTCGAGACTCAACG 2975  
 QY 774 yLeuIleThrSerAspTyrGlyLeuLeuSerAspProthIleProGlyIleProSerProAr 794  
 DB 2976 CCTCATTTACATCAGATTAGAGAGCTCTCTCA-GACCCACACACCTGGGATCCCGAGTCTCG 3034  
 QY 794 gAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGly 814  
 DB 3035 AGATGAGCTGTGG--T--GGCCCGACACTATATGCTCTGCGAGAGACCCCGCCATATC-CA 3089  
 QY 814 uGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValIle 834  
 DB 3090 GGAGAGACACACTTAAGTACATCTCTTGTGGGCAAGGGCACTTTGGCAGCTGGAGCT 3149  
 QY 834 uCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGly 854

DB 3150 GTGCCGCTATGACCC--TG-GACAAATACGGGACCCCTGTGGCTGAGTGAACAGCTACA 3206  
 QY 854 nHisSerVal-ProAspGlnGlnArgAspPheGlnArgGlnIleGlnIleLeuLysAlaL 874  
 DB 3207 GCACAGCGG-GCCAGACCCAGCAGAGGACTTCCAGCGGAGATTGATGATCTTAAGGCTC 3265  
 QY 874 euhHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerL 894  
 DB 3266 TGCACAGCGACTCATCGTCAAGTACCGGGGATGACGATAGGCGCAGAGTCCGACAGCGC 3325  
 QY 894 euArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeu-LeuGlnArgHis 913  
 DB 3326 TGGCGTTGGTGAAGAGTACTTCCAGCGGCTGCTGGAGAG-CTTCTCTGACGCCCAT 3384  
 QY 914 --ArgGly-LeuHisThrAspArgLeuLeuPheAlaTyrPheIleCysLysGlyMet 932  
 DB 3385 CGCGCGCCCTGCACACCGACCGCTACTGCTGTCTTGGCTTGGCAGATCTGCAAGGCGATG 3444  
 QY 933 GluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 DB 3445 GAGTACTGTGGTGGCGCGCTGCTGACACCGTGAACCTGGCTGGCGCAACATCTTGTTG 3504  
 QY 953 GluSerGlnAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGly 972  
 DB 3505 GAGAGCGAGGCTCATGTGAAGATCGCGGACTTGGCTGCTGAAGCTGTGCCCCCTGGGA 3564  
 QY 973 LysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGly 992  
 DB 3565 AAGGACTACTAGCTGTCCCGAGCGCTGGCCAAAGCCCATCTTTTGGTATGCGCCGAG 3624  
 QY 993 SerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeu 1012  
 DB 3625 TCCCTATCTGACAAATCTTCTCCCGCAATCTGAGTGTGGAGCTTGGAGTGGTGTG 3684  
 QY 1013 TyrGluLeuPheThrTyrCysAspLysSerCysSerProSerIleGluPheLeuArgMet 1032  
 DB 3685 TACGAGCTCTTCACTAGCTGCGACAGAGCTGAGCGCCATCCCTGATGCTTGGCGATG 3744  
 QY 1033 MetGlyProGluArgGlyGlyProProLeuCysArgLeuLeuGluLeuLeuAlaGlyGly 1052  
 DB 3745 ATGGGGCTGAGGCTGAGAGACCCCGCTCTGCTGAGTGGTGGAGAGGCG 3804  
 QY 1053 ArgArgLeuProProProthIleCysProthIleGluValGlnLeuMetGlnLeuCys 1072  
 DB 3805 CGAGCGCTCCACACACTCCACCTGCCACCGAGGTTTCAGAGAGCTCATGACAGTGTGC 3864  
 RESULT 2  
 AAT30862  
 ID AAT30862 standard; DNA; 3807 BP.  
 AC AAT30862;  
 XX  
 DT 13-SEP-1996 (first entry)  
 XX  
 DE Protein tyrosine kinase JAK3 cDNA.  
 XX  
 DE JAK3: protein tyrosine kinase; cell proliferation; differentiation;  
 KW signal transduction; leukaemia; aplastic anaemia; myelodysplasia;  
 KW polycythemia vera; thrombocytosis; gene therapy; diagnosis; ss.  
 OS Homo sapiens.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 168..3359  
 FT /\*tag= a  
 FT polyA\_signal 3692..3696  
 FT /\*tag= b  
 FT polyA\_signal 3793..3797  
 FT /\*tag= c  
 FT PN W09618639-A1.

PD 20-JUN-1996.  
 XX  
 PF 15-DEC-1995; 95WO-US16435.  
 XX  
 PR 15-DEC-1994; 94US-0357598.  
 XX  
 PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MED.  
 XX  
 PI Clavin CI, Safford MG, Small D;  
 XX  
 DR WPI: 1996-300568/30.  
 DR P-PSDB: AAR96037.  
 XX  
 PT Protein tyrosine kinase, JAK3, protein and nucleic acid - used in  
 PT the gene therapy of cellular proliferative diseases, e.g. leukaemia,  
 PT aplastic anaemia etc.  
 XX  
 PS Claim 4; Page 43-47; 97pp; English.  
 XX  
 CC A cDNA clone (AAT30862) codes for JAK3 (AAR96037), a new member of the  
 CC JAK family of non-receptor protein tyrosine kinases, that probably  
 CC plays a role in growth factor modulated differentiation,  
 CC proliferation and survival of haematopoietic stem/progenitor cells.  
 CC It was derived from CD34+ mRNA obtd. from normal human bone marrow  
 CC by PCR amplification using primers (see also AAT30863-64) based  
 CC on highly conserved motifs from protein tyrosine kinase catalytic  
 CC domains. The JAK3 coding sequence can be incorporated into a  
 CC vector and used for prodn. of recombinant JAK3. It can also be  
 CC used in gene therapy protocols for leukaemia, myelodysplasia,  
 CC polthemia vera, thrombocytosis and aplastic anaemia, or to  
 CC stimulate haematopoietic cell proliferation. The gene was  
 CC localised to chromosome 19, band p12-13.1.  
 XX  
 SQ Sequence 3807 BP; 716 A; 1219 C; 1099 G; 773 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. NO.: 2.37e-48 Length: 3807  
 Score: 62.00 Matches: 96  
 Percent Similarity: 97.96% Conservative: 0  
 Best Local Similarity: 97.96% Mismatches: 1  
 Query Match: 5.64% Indels: 2  
 DB: 17 Gaps: 0  
 US-09-397-967-16 (1-1099) x AAT30862 (1-3807)  
 QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 2985 CGCGGTCGCGGACACCGGAGACTGCGCCCGGAAACATCCTCGTGAGAGAGGAGGACAC 3044  
 QY 958 ValIysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 3045 GTCAAGATCGCTGACTTCGCGCTAGCTAGCTGCTGCGCGCTTGAC-AAAGACTACTAGCT 3103  
 QY 977 lValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAs 997  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 3104 GGTCCGGAGGACGACGACCCCATTTCTGTGATGCCGCCGAATCCCTCTGGAGACAA 3163  
 QY 997 nIlePheSerArgGlnSerAspValTyrPheGlyValValLeuTyrGluLeuPheTh 1017  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 3164 CATTTCTCTGCGCAGTCAGACGTCGTGAGACTTCGGGGTCTCTGTACGAGCTCTTAC 3223  
 QY 1017 TTYrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 3224 CTACTGCGACAAAGCTGACGCCCTCGCGCGAGTCTCTGCGGATGGGA 3275  
 RESULT 3  
 AAS10808  
 1D AAS10808 standard; cDNA; 795 BP.  
 XX  
 AC AAS10808;  
 XX  
 XX 24-~~Oct~~-2001 (first entry)  
 XX

DE Human Janus kinase 3 (JAK3) partial cDNA, 3-2 primer complement.  
 XX  
 KW Human; Janus kinase 3; JAK3; ss; JAK/STAT inhibitor; 3-2 primer;  
 KW signal transducer and activator of transcription; osteoarthritis;  
 KW degenerative joint disease; rheumatoid arthritis; leprosy; asthma;  
 KW cancer; tumour; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200152892-A2.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 22-JAN-2001; 2001WO-US02033.  
 XX  
 PR 24-JAN-2000; 2000US-0177872.  
 PR 28-NOV-2000; 2000US-0723490.  
 XX  
 PA (GENZ ) GENZYME CORP.  
 XX  
 PI Vastios G;  
 XX  
 DR WPI: 2001-465338/50.  
 XX  
 PT Use of inhibitors of Janus kinase/signal transducers and activators of  
 PT transcription for inhibiting onset and progression of degenerative  
 PT joint diseases or disorders such as osteoarthritis, rheumatoid  
 PT arthritis  
 XX  
 PS Example 4; Fig 3; 55pp; English.  
 XX  
 CC The sequence is a partial sequence, designated "complement of  
 CC clone 3-2 primer", obtained from a chondrocyte cDNA encoding  
 CC human Janus kinase 3 (JAK3). The invention relates to the use of  
 CC JAK/STAT (Janus kinase/signal transducer and activator of  
 CC transcription) inhibitors other than debromymentadistine (DBM) and  
 CC hymenaldistine (H) for inhibiting the progression or the likelihood of  
 CC developing diseases involving cartilage degradation, and for regulating  
 CC the expression of pro-inflammatory agents or cytokines in a chondrocyte,  
 CC and cartilage degrading enzymes in a cell. A JAK3/STAT inhibitor of the  
 CC invention is useful for inhibiting progression or likelihood of  
 CC developing osteoarthritis or rheumatoid arthritis. The inhibitor is also  
 CC useful for treating other JAK/STAT-mediated diseases or disorders,  
 CC including T cell-mediated disorders, mast cell-mediated disorders,  
 CC type 2 (cytokine hypersensitivity) disorders, B cell lymphoma, and  
 CC myeloid diseases. T cell-mediated disorders include human T cell  
 CC leukaemia/lymphoma virus (HTLV)-1, Sdzory's syndrome, c-rabl  
 CC transformation, natural killer-like T cell lymphomas (NK-like tumours)  
 CC and graft-vs-host disease; cytokine hypersensitivity disorders include  
 CC leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated  
 CC disorders include hay fever, asthma, hives and anaphylaxis; and  
 CC leukaemias and lymphomas include acute lymphocytic and lymphoblastic  
 CC leukaemias, B cell lymphomas and leukaemias of myeloid origin. DBM and H  
 CC are useful as therapeutic agents in cancers in which JAK3 plays a role  
 CC in the initiation or progression of tumorigenesis.  
 XX  
 SQ Sequence 795 BP; 156 A; 255 C; 226 G; 158 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. NO.: 1.55e-38 Length: 795  
 Score: 51.00 Matches: 51  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.64% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-397-967-16 (1-1099) x AAS10808 (1-795)  
 QY 858 ProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAsp 877  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 455 CCAGACGACGACGAGGAGACTTTCACGCGGAGATTCAGATCCCAAGACGTCGACAGTGAT 514  
 QY 878 PheIleValIysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuVal 897

```

Cc Db 515 TTCAATGTCAGATATCGTGTCTAGTATGCGCGCGCCGACAGCCTGGCTGTC 574
Oy 898 MetGluTyrLeuProSerGlyCysLeuArgasp 908
Db 575 ATGGAGTACTGCGCCAGCGGCTGCTGGCGCAGC 607

RESULT 4
AA010809
ID AA010809 standard: cDNA: 778 BP.
XX
XX AA010809:
AC
XX
XX 24-OCR-2001 (first entry)
DT
XX
XX Human Janus kinase 3 (JAK3) partial cDNA sequence, 3-2 primer 1.
DE
XX
XX Human; Janus kinase 3; JAK3; ss; JAK/STAT inhibitor; 3-2 primer 1;
KW signal transducer and activator of transcription; osteoarthritis;
KW degenerative joint disease; rheumatoid arthritis; leprosy; asthma;
KW cancer; tumour; leukaemia.
XX
XX Homo sapiens.
OS
XX
XX WO200152892-A2.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 22-JAN-2001; 2001WO-US02033.
PE
XX
XX 24-JAN-2000; 2000US-0177872.
PR
XX 28-NOV-2000; 2000US-0723490.
PA (GENZ ) GENZYME CORP.
XX
XX
XX Vastios G;
PI
XX
XX WPI: 2001-465338/50.
DR
XX
XX
XX Use of inhibitors of Janus kinase/signal transducers and activators of
PT transcription for inhibiting onset and progression of degenerative
PT joint diseases or disorders such as osteoarthritis, rheumatoid
PT arthritis
XX
XX
XX Example 4; Fig 3; 55pp; English.
PS
XX
XX The sequence is a partial sequence, designated "clone 3-2 primer 1",
CC obtained from a chondrocyte cDNA encoding human Janus kinase
CC 3 (JAK3). The invention relates to the use of JAK/STAT (Janus
CC kinase/signal transducer and activator of transcription)
CC inhibitors other than debromohymenialdisine (DBH) and
CC hymenialdisine (H) for inhibiting the progression or the likelihood of
CC developing diseases involving cartilage degradation, and for regulating
CC the expression of pro-inflammatory agents or cytokines in a chondrocyte,
CC and cartilage degrading enzymes in a cell. A JAK3/STAT inhibitor of the
CC invention is useful for inhibiting progression or likelihood of
CC developing osteoarthritis or rheumatoid arthritis. The inhibitor is also
CC useful for treating other JAK/STAT-mediated diseases or disorders,
CC including T cell-mediated disorders, mast cell-mediated disorders,
CC type 2 (cytokine hypersensitivity) disorders, B cell lymphoma, and
CC myeloid diseases. T cell-mediated disorders include human T cell
CC leukaemia/lymphoma virus (HTLV)-1, Sclar's syndrome, c-abl
CC transfection, natural killer-like T cell lymphomas (NK-like tumours)
CC and graft-vs-host disease; cytokine hypersensitivity disorders include
CC leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated
CC disorders include hay fever, asthma, hives and anaphylaxis; and
CC leukaemias and lymphomas include acute lymphocytic and lymphoblastic
CC leukaemias. B cell lymphomas and leukemias of myeloid origin. DBH and H
CC are useful as therapeutic agents in cancers in which JAK3 plays a role
CC in the initiation or progression of tumorigenesis.
XX
XX Sequence 778 BP; 131 A; 264 C; 220 G; 163 T; 0 other;

```

```

Alignment Scores:
Pred. No.: 4,33e-28 Length: 778
Score: 40.00 Matches: 74
Percent Similarity: 97.37% Conservative: 0
Best Local Similarity: 97.37% Mismatches: 1
Query Match: 3,64% Indels: 2
DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x AA010809 (1-778)
Oy 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957
Db 103 CGCCGCTGCGTGCACCGGACCTGGCCGCCAACAATCTCTGTGAGAGCGAGCGACAC 162
Oy 958 ValIysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977
Db 163 GTCAAGATCGCTGACTTGGCTTGGCTTACTAGCTGCTGCCGCTTGAC-AAAGACTACTAGCT 221
Oy 977 lValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAs 997
Db 222 GGTCCGCGAGCCAGCGCCAGAGCCCATTTTCTGTATGCCCCGAAATCCCTCTCGACAA 281
Oy 997 nIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeu 1012
Db 282 CATCTTCTCTCGCAGTCAAGACGTCTGAGCTTGCGGGGTCTCTG 327

RESULT 5
AA087512
ID AA087512 standard: cDNA to mRNA; 165 BP.
XX
XX AA087512:
AC
XX
XX 22-NOV-1995 (first entry)
DT
XX
XX Tyrosine kinase peptide coding sequence.
DE
XX
XX Tyrosine kinase peptide; UT-7 cell; primer: polymerase chain reaction;
KW PCR; cell differentiation; antisense drug; ds.
XX
XX
XX Homo sapiens.
OS
XX
XX JP07059569-A.
PN
XX
XX 07-MAR-1995.
PD
XX
XX 25-AUG-1993; 93JP-0210404.
PE
XX
XX 25-AUG-1993; 93JP-0210404.
PR
XX
XX (ASAH ) ASAH KASEI KOGYO KK.
PA
XX
XX WPI: 1995-135894/18.
DR
XX P-PSDB: AAR71394.
XX
XX New DNA sequence encoding a tyrosine kinase peptide - useful for
PT evaluation and control of cell differentiation.
PT
XX
XX Claim 1; Page 5; 5pp; Japanese.
PS
XX
XX This sequence encodes a tyrosine kinase peptide which is derived
CC from UT-7 cells. This sequence was isolated using the primer
CC sequences given in AA087510-11. The tyrosine kinase peptide may be
CC used for the evaluation and control of cell differentiation and is
CC also useful for the development of drugs and antisense drugs.
XX
XX
XX Sequence 165 BP; 36 A; 56 C; 39 G; 34 T; 0 other;

```

```

Alignment Scores:
Pred. No.: 3,61e-20 Length: 165
Score: 31.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2,82% Indels: 0

```

DB: 16 Gaps: 0

US-09-397-967-16 (1-1099) x AAG87512 (1-165)

OY 973 LysAspTyrTyrValValArgIleuProGlyGlnSerProIlePheTrpTyrAlaProGlu 992  
 |||  
 DB 73 AAGAGCTACTGAGTGTGTCGGCAGCCAGCCAGACCCCATTTTGTGATGCCCCCGAA 132  
 |||  
 OY 993 SerLeuSerAspAsnIlePheSerArgGlnSer 1003  
 |||  
 DB 133 TCCCTCTCGGACAAACATCTCTCTCGCAGTCA 165  
 |||

RESULT 6  
 AAS10806  
 ID AAS10806 standard; cDNA: 811 BP.  
 AC AAS10806:  
 XX  
 XX 24-OCT-2001 (first entry)  
 DE Human Janus kinase 3 (JAK3) partial cDNA sequence, 3-2 KS.  
 XX  
 KW Human; Janus kinase 3; JAK3; ss; JAK/STAT inhibitor; 3-2 KS;  
 KW signal transducer and activator of transcription; osteoarthritis;  
 KW degenerative joint disease; rheumatoid arthritis; leprosy; asthma;  
 KW cancer; tumour; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200152892-A2.  
 PD 26-JUL-2001.  
 XX  
 XX 22-JAN-2001; 2001WO-US02033.  
 PF  
 XX 24-JAN-2000; 2000US-0177872.  
 PR 28-NOV-2000; 2000US-0723490.  
 PA (GEN2 ) GENZYME CORP.  
 PI  
 PI Vastios G;  
 XX  
 XX WPI; 2001-465338/50.  
 DR  
 XX  
 XX Use of inhibitors of Janus kinase/signal transducers and activators of  
 PT transcription for inhibiting onset and progression of degenerative  
 PT joint diseases or disorders such as osteoarthritis, rheumatoid  
 PT arthritis -  
 XX  
 PS Example 4; Fig 3; 55pp; English.

The sequence is a partial sequence from a chondrocyte cDNA encoding  
 CC human Janus kinase 3 (JAK3). The invention relates to the use of  
 CC JAK/STAT (Janus kinase/signal transducer and activator of  
 CC transcription) inhibitors other than debromohymenialdisine (DBH) and  
 CC hymenialdisine (H) for inhibiting the progression of the likelihood of  
 CC developing diseases involving cartilage degradation, and for regulating  
 CC the expression of pro-inflammatory agents or cytokines in a chondrocyte,  
 CC and cartilage degrading enzymes in a cell. A JAK3/STAT inhibitor of the  
 CC invention is useful for inhibiting progression or likelihood of the  
 CC developing osteoarthritis or rheumatoid arthritis. The inhibitor is also  
 CC useful for treating other JAK/STAT-mediated diseases or disorders,  
 CC including T cell-mediated disorders, mast cell-mediated disorders,  
 CC type 2 (cytokine hypersensitivity) disorders, mast cell lymphoma, and  
 CC myeloid diseases. T cell-mediated disorders include human T cell  
 CC leukemia/lymphoma virus (HTLV)-1, Sdcory's syndrome, c-abl  
 CC transformation, natural killer-like T cell lymphomas (NK-like tumours)  
 CC and graft-vs-host disease; cytokine hypersensitivity disorders include  
 CC leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated  
 CC disorders include hay fever, asthma, hives and anaphylaxis; and  
 CC leukemias and lymphomas include acute lymphocytic and lymphoblastic  
 CC leukemias, B cell lymphomas and leukaemias of myeloid origin. DBH and H  
 CC are useful as therapeutic agents in cancers in which JAK3 plays a role

CC in the initiation or progression of tumourigenesis.  
 XX  
 XX Sequence 811 BP; 149 A; 254 C; 226 G; 182 T; 0 other;  
 SQ

Alignment Scores:

Pred. No.:	Length:	811
Score:	29.00	Matches: 29
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	2.64%	Indels: 0
DB:	22	Gaps: 0

US-09-397-967-16 (1-1099) x AAS10806 (1-811)

OY 1006 TrpSerPheGlyValValIleuTyrGluIleuPheTrpTyrCysAspLysSerCysSerPro 1025  
 |||  
 DB 168 TGGACCTTCGGGGTCTCTCTTACGAGCTCTTACCTACTGCGACAAAGCTGACGCC 227  
 |||

OY 1026 SerAlaGluPheLeuValArgMetMetGly 1034  
 |||  
 DB 228 TCGGCCGAGTTCCTCGCATGATGGGA 254  
 |||

RESULT 7  
 ABN21423  
 ID ABN21423 standard; cDNA: 347 BP.  
 XX  
 XX ABN21423:  
 AC  
 AC  
 XX 24-JUN-2002 (first entry)  
 DE  
 DE Human ORFX polynucleotide sequence SEQ ID NO:11323.  
 XX  
 XX  
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
 KW hypertension; hypothyroidism; cholesterol ester storage disease;  
 KW immune deficiency; immune disorder; infectious disease;  
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
 KW myasthenia gravis; gene; ss.  
 XX  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200192523-A2.  
 PN  
 PN  
 XX 06-DEC-2001.  
 PD  
 PD  
 XX 29-MAY-2001; 2001WO-US10836.  
 PF  
 XX 30-MAY-2000; 2000US-206132P.  
 PR 29-AUG-2000; 2000US-228716P.  
 PA (CURA-) CURAGEN CORP.  
 PI  
 PI Shimkets RA, Leach MD;  
 XX  
 XX WPI; 2002-106308/14.  
 DR P-PSDB; ABP05671.  
 XX

Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and autoimmune disorders -  
 XX

Disclosure; SEQ ID 11323; 1037pp; English.

The present invention describes substantially purified human proteins  
 CC (referred to as open reading frame, ORFX, where x is 1-11491 (see Table 1  
 CC in the specification)). ABN15762 to ABN27252 encode the human ORFX  
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for  
 CC treating or preventing a pathology associated with an ORFX-associated  
 CC disorder in humans, and in the manufacture of a medicament for treating a  
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
 CC sequences can be used in gene therapy. ORFX sequences can be used in the



CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
CC lupus erythematosus, hypertension, hypochylodism, cholesterol ester  
CC storage disease, various immune deficiencies and disorders, infectious  
CC diseases, autoimmune thyroiditis, myasthenia gravis, rheumatoid  
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
CC disease and autoimmune inflammatory eye disease. ORF proteins are also  
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
CC bone degenerative disorders, or periodontal disease, and for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues and conditions resulting from  
CC systemic cytokine damage.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
XX  
SO Sequence 347 BP; 67 A; 117 C; 90 G; 69 T; 4 other;

Alignment Scores:  
Pred. No.: 4,59e-16 Length: 347  
Score: 27.00 Matches: 27  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2,46% Indels: 0  
DB: Gaps: 24

US-09-397-967-16 (1-1099) x ABN21423 (1-347)

Oy 1043 CysArgLeuLeuGluLeuAlaGluGlyArgArgLeuProProProProProPro 1062  
Db 19 TCGCCCTCTCGAGCTCTGGCAGAGGCGCGCTCCACCTCCACCTGCC 78

Oy 1063 ThrGluValGluGluLeuMet 1069  
Db 79 ACCGAGGTTCAGGACGCTCATG 99

RESULT 8  
AAK78095  
ID AAK78095 standard; DNA; 3222 BP.  
XX  
AC AAK78095;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32907.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225269.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 01-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231245.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0233397.  
PR 14-SEP-2000; 2000US-0233398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0233400.  
PR 14-SEP-2000; 2000US-0233401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-024617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2001US-0259678.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-483426/52.  
 DR  
 XX  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 XX metastasis -  
 XX  
 PS Disclosure; SEQ ID NO 32907; 3071pp + Sequence Listing; English.  
 XX  
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patients own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the

CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 3222 BP; 721 A; 883 C; 895 G; 723 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,35e-11 Length: 3222  
 Score: 23.00 Matches: 23  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.09% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-397-967-16 (1-1099) x AAK78095 (1-3222)  
 QY 325 GUAAGlupheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyr 344  
 DB 380 GAGCCGAGTTCCCGAGGCTGCTGTGCTTCGTGGCGCTGTGAGCGCTAC 439  
 QY 345 PheArgLeu 347  
 DB 440 TTCCGCTG 448  
 RESULT 9  
 ID AAK57677 standard; DNA; 65 BP.  
 XX  
 AC AAK57677;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:30425.  
 XX  
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;  
 KW splice variant; transcriptome; oligonucleotide library; ss.  
 XX  
 OS Mus musculus.  
 XX  
 PN W0200210449-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 20-JUL-2001; 2001WO-IB01903.  
 XX  
 PR 28-JUL-2000; 2000US-221607P.  
 PR 02-MAY-2001; 2001US-287724P.  
 PA (COMP-) COMPUGEN INC.  
 XX  
 PI Shoshan A, Wasserman A, Mintz E, Mintz L, Falgier S;  
 XX  
 DR WPI; 2002-257383/30.  
 XX  
 PT New oligonucleotide libraries comprising oligonucleotides which  
 PT selectively hybridize to mRNAs transcribed from a transcription unit of  
 PT a genome, useful for detecting tissue-, pathology-, and  
 PT developmental-specific genes -  
 XX  
 PS Example 1; SEQ ID 30425; 47pp; English.  
 XX  
 CC The present invention describes oligonucleotide libraries for detecting  
 CC messenger RNAs that populate a (sub-)transcriptome, where the  
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
 CC transcription units that populate a genome. The library comprises  
 CC several oligonucleotides, each capable of hybridizing selectively to a  
 CC set of messenger RNAs transcribed from a given transcription unit of  
 CC the genome, which encodes one or more messenger RNA splice variants.  
 CC The oligonucleotide libraries are useful for detecting mRNAs from a  
 CC biological sample, in expression profiling studies, in qualitatively or

CC quantitatively characterizing the corresponding transcriptome, and in  
 CC detecting RNA transcripts and splice variants of human or animal  
 CC transcriptomes. The libraries may also be used as specialised mini  
 CC libraries to detect transcripts of a sub-transcriptome under a  
 CC particular biological or pathological state, and so allowing the  
 CC detection of tissue- and pathology-specific genes such as those genes  
 CC only expressed in specific tissue under a specific pathological  
 CC condition; to detect developmental specific genes; and to detect RNA  
 CC transcripts and splice variants of a transcriptome of a patient suffering  
 CC from a particular disorder. ABN27253 to ABN39589 represent  
 CC oligonucleotide sequences from rats, humans and mice, which are used in  
 CC the exemplification of the present invention.  
 CC N.B. The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 65 BP; 16 A; 18 C; 17 G; 14 T; 0 other;

#### Alignment Scores:

Pred. No.: 4,81e-11 Length: 65  
 Score: 21.00 Matches: 21  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.91% Indels: 0  
 DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x ABN57677 (1-65)

OY 739 LeuProAlaLeuysTrpRhgIuLeuAlaGlyLeuIllethrgIncysMetAlaTyasp 758

DB 3 CTGGCCCTCTCAATGGACAGAACTGGCGACTTATCACACAGTGCATGGCGTATGAT 62

OY 759 Pro 759

DB 63 CCT 65

RESULT 10

AA510804  
 ID AAS10804 standard; cDNA: 109 BP.

AC AAS10804;

DT 24-OCT-2001 (first entry)

DE Human Janus kinase 3 (JAK3) RT-PCR fragment #1.

KW Human; Janus kinase 3; JAK3; RT-PCR; ss: JAK/STAT inhibitor;

KW signal transducer and activator of transcription; osteoarthritis;

KW degenerative joint disease; rheumatoid arthritis; leprosy; asthma;

KW cancer; tumour; leukaemia.

OS Homo sapiens.

PN WO200152892-A2.

PD 26-JUL-2001.

PF 22-JAN-2001; 2001WO-US02033.

PR 24-JAN-2000; 2000US-0177872.

PR 28-NOV-2000; 2000US-0723490.

PA (GENZ ) GENZYME CORP.

PI Vastios G;

DR WPI; 2001-465338/50.

Use of inhibitors of Janus kinase/signal transducers and activators of  
 transcription for inhibiting onset and progression of degenerative  
 joint diseases or disorders such as osteoarthritis, rheumatoid  
 arthritis

PS Example 2; Fig 1; 55pp; English.

XX The sequence is obtained from an RT (reverse transcriptase)-PCR fragment  
 CC from human Janus kinase 3 (JAK3). The invention relates to the use of  
 CC JAK/STAT (Janus kinase/signal transducer and activator of  
 CC transcription) inhibitors other than dephosphorylating (DBP) and  
 CC hmyanialdisine (H) for inhibiting the progression or the likelihood of  
 CC the expression of pro-inflammatory agents or cytokines in a chondrocyte,  
 CC and cartilage degrading enzymes in a cell. A JAK3/STAT inhibitor of the  
 CC invention is useful for inhibiting progression or likelihood of  
 CC developing osteoarthritis or rheumatoid arthritis. The inhibitor is also  
 CC useful for treating other JAK/STAT-mediated diseases or disorders,  
 CC including T cell-mediated disorders, mast cell-mediated disorders,  
 CC type 2 (cytokine hypersensitivity) disorders, B cell lymphoma, and  
 CC myeloid diseases. T cell-mediated disorders include human T cell  
 CC leukaemia/lymphoma virus (HTLV)-1, Sotor's syndrome, C-abl  
 CC transformation, natural killer-like T cell lymphomas (NK-like tumours)  
 CC and graft-vs-host disease; cytokine hypersensitivity disorders include  
 CC leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated  
 CC disorders include hay fever, asthma, hives and anaphylaxis; and  
 CC leukaemias, B cell lymphomas include acute lymphocytic and lymphoblastic  
 CC leukaemias, B cell lymphomas and leukaemias of myeloid origin. DBP and H  
 CC are useful as therapeutic agents in cancers in which JAK3 plays a role  
 CC in the initiation or progression of tumorigenesis.

SO Sequence 109 BP; 22 A; 36 C; 34 G; 15 T; 2 other;

#### Alignment Scores:

Pred. No.: 5.54e-08 Length: 109  
 Score: 18.00 Matches: 18  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.64% Indels: 0  
 DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x AAS10804 (1-109)

OY 945 LeuAlaAlaArgaenilleuValGluSerGluAlaHisValIysIleAlaasp 962

DB 56 TTGGCCGCCGAAACATCTCTGGAGAGCGAGCAGCAGCTCAATCGCTGAC 109

RESULT 11

AA510807  
 ID AAS10807 standard; cDNA: 726 BP.

AC AAS10807;

DT 24-OCT-2001 (first entry)

DE Human Janus kinase 3 (JAK3) partial cDNA sequence, 3-2 SK.

KW Human; Janus kinase 3; JAK3; ss: JAK/STAT inhibitor; 3-2 SK;

KW signal transducer and activator of transcription; osteoarthritis;

KW degenerative joint disease; rheumatoid arthritis; leprosy; asthma;

KW cancer; tumour; leukaemia.

OS Homo sapiens.

PN WO200152892-A2.

PD 26-JUL-2001.

PF 22-JAN-2001; 2001WO-US02033.

PR 24-JAN-2000; 2000US-0177872.

PR 28-NOV-2000; 2000US-0723490.

PA (GENZ ) GENZYME CORP.

PI Vastios G;

DR WPI; 2001-465338/50.

7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

[illegible]

PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 01-NOV-2000; 2000US-0241826.  
 PR 08-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 17-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249246.  
 PR 17-NOV-2000; 2000US-0249247.  
 PR 17-NOV-2000; 2000US-0249248.  
 PR 17-NOV-2000; 2000US-0249249.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 08-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-02559678.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-483426/52.  
 XX  
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -  
 XX  
 PS Disclosure: SEQ ID NO 32909; 3071bp + Sequence Listing; English.  
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)

CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patients own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/haematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/haematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 2000 BP; 345 A; 609 C; 640 G; 406 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 8,49e-07 Length: 2000  
 Score: 18.00 Matches: 18  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.64% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-397-967-16 (1-1099) x AAK78097 (1-2000)  
 QY 1 MetAlaProPserGlnGluThrProLeuIleProGlnArgSerCysSerIleu 18  
 Db 25 ATGGCAGCTCCAGTGAAGAGAGCGCCCTGATCCCTCAGCGTTATCAGCGCTC 78  
 RESULT 13  
 AAV61801  
 ID AAV61801 standard; cDNA: 3435 BP.  
 XX  
 XX AAV61801;  
 AC 20-JAN-1999 (first entry)  
 XX  
 DT 20-JAN-1999 (first entry)  
 XX  
 DE JAK2 protein encoding cDNA.  
 XX  
 KW JAK2 protein; cytoplasmic domain; beta subunit; screening; asthma;  
 KW Interleukin; granulocyte macrophage-colony stimulating factor; GM-CSF;  
 KM IL-3; IL-5; human; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH  
 FT Key Location/Qualifiers  
 FT CDS 10..3429 /\*tag= a  
 FT /\*product= "JAK2 protein"  
 XX  
 PN WO9843087-A1.  
 XX  
 PD 01-OCT-1998.  
 XX  
 XX 23-MAR-1998; 98MO-US05387.  
 XX  
 PR 24-MAR-1997; 97US-0041511.  
 XX  
 PA (PHAA) PHARMACIA & UPJOHN CO.  
 XX  
 XX Kayles PS, Roberds SL;  
 PI WPI; 1998-532151/45.  
 XX  
 DR P-PSDB; AAM76425.  
 XX  
 PT Screening for compounds useful for preventing or treating asthma -  
 PT by determining if compounds inhibit binding of the JAK2 protein to  
 XX e.g. IL-3, IL-5 or GM-CSF

PS Disclosure: Fig 22A-C; 112pp; English.

CC This cDNA encodes the full-length JAK2 protein. The N-terminal fragment of JAK2 protein can be used in the method of the invention of screening for compounds useful for treating or preventing asthma. The method 294 comprises contacting a molecule comprising at least the N-terminal 294 amino acid residues of the JAK2 protein, with another molecule comprising at least 13 membrane-proximal cytoplasmic amino acids of Interleukin (IL)-3, IL-5 or granulocyte macrophage-colony stimulating factor (GM-CSF) proteins in the presence of the candidate compound, and determining whether the first and the second molecules form a complex. If the compound inhibits complex formation, it can be used to treat asthma.

XX Sequence 3435 BP; 1157 A; 598 C; 733 G; 947 T; 0 other;

SO Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.000112	3435	16	0	0	0
Percent Similarity:	100.00%					
Best Local Similarity:	100.00%					
Query Match:	1.46%					

DB: 19

US-09-397-967-16 (1-1099) x AAV61801 (1-3435)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
 |||  
 DB 3106 TCAGATGTGTGGAGCTTGGAGTGTCTGTATGAACCTTTTCACATAC 3153

RESULT 14

ID AAQ25307 standard; DNA: 3473 BP.

XX AAQ25307:

AC 18-FEB-1999 (first entry)

DT JAK2 encoding DNA.

DE JAK2 encoding DNA.

XX Phosphorylation: JAK1, JAK2; protein tyrosine kinase; human;

KW catalytic domain; SH2 domain; growth factor receptor; PTK; murine; ss.

XX Mus musculus.

OS Mus musculus.

XX Key 1.2961 Location/Qualifiers

FT CDS /tag= a

FT /product= JAK2

XX MO9210519-A.

PN 25-JUN-1992.

PD 26-NOV-1991; 91WO-US08889.

PF 28-NOV-1990; 90AU-0003594.

PR (LUDWIG) INST CANCER RES.

XX (LUDWIG) INST CANCER RES.

PA Harpur A, Wilks AF, Ziemiecki A;

PI WPI: 1992-234591/28.

DR P-PSDB; AAR25141.

PT Novel protein tyrosine kinase mol. - comprises multiple catalytic domains but no SH2 domain and is for phosphorylation of proteins

XX domains but no SH2 domain and is for phosphorylation of proteins

PS Claim 10; Fig 8; 50pp; English.

XX This sequence encodes the murine protein tyrosine kinase JAK2 (from CC Janus kinase). Northern analysis of JAK2 expression in a mouse CC demonstrated the presence of two mRNA transcripts (4.8 and 4.4 kb) CC The levels of these transcripts alter with respect to one another in

CC different tissues. The kidney, spleen and lung appear to express CC predominantly the larger form, whereas ovary, placenta, skeletal muscle CC and all murine cell lines analysed express both forms at equal levels. CC The difference in sizes may be due to differential polyadenylation CC sites. Both JAK2 and JAK1 are examples of a new subfamily or class CC of protein tyrosine kinase. These can be used in the phosphorylation CC of proteins, incorporation of labels and in the design of analogues, CC antagonists and agonists of JAK's.

XX Sequence 3473 BP; 1128 A; 677 C; 781 G; 887 T; 0 other;

SO Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.000113	3473	16	0	0	0
Percent Similarity:	100.00%					
Best Local Similarity:	100.00%					
Query Match:	1.46%					

DB: 13

US-09-397-967-16 (1-1099) x AAQ25307 (1-3473)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
 |||  
 DB 2674 TCAGATGTGTGGAGCTTGGAGTGTCTGTATGAACCTTTTCACATAC 2721

RESULT 15

ID AAD22680 standard; cDNA: 3495 BP.

XX AAD22680:

AC 26-FEB-2002 (first entry)

DT Murine JAK2 tyrosine kinase cDNA.

DE Murine JAK2 tyrosine kinase cDNA.

XX Murine; Src homology 2-Bbeta; SH2-Bbeta; neuroprotective; gene therapy; cell differentiation; nerve regeneration; angiogenesis; embryogenesis; CC cytosolic; antisense therapy; drug screening; cellular expression; CC immunological disease; neurological disease; apoptosis; diabetes; cancer; arthritis; JAK2 tyrosine kinase; ss.

XX Mus sp.

OS Mus sp.

XX Key 1.2982 Location/Qualifiers

FT CDS /tag= a

FT /product= "Murine JAK2 tyrosine kinase protein"

XX US6312941-B1.

PN 06-NOV-2001.

PD 26-NOV-1997; 97US-0980080.

PF 26-NOV-1997; 97US-0980080.

PR (UNMI) UNIV MICHIGAN.

XX (UNMI) UNIV MICHIGAN.

PA Carter-Su C, Rui L, Karow DS;

PI WPI: 2002-024907/03.

DR P-PSDB; AAE13587.

PT New isolated nucleic acid molecule encoding full length Src homology 2 (SH2)-Bbeta protein, useful for modulating cellular expression of active SH2-Bbeta protein

XX Example 1; Fig 3A; 71pp; English.

XX The invention relates to human Src homology 2 (SH2)-Bbeta protein and its CC DNA molecule. SH2-Bbeta is useful in drug screening assays designed to CC identify drugs that interfere with the specific binding of JAK2 tyrosine CC kinase which is a member of Janus family of tyrosine kinase-interacting

CC signaling molecule. SH2-Bbeta protein is useful for modulating,  
 CC preferably reducing cellular expression or intracellular concentration or  
 CC availability of active SH2-Bbeta. SH2-Bbeta is useful for treating a  
 CC disease associated with undesirable cell growth, differentiation, and  
 CC growth-factor/cytokine, preferably interleukin, more preferably growth  
 CC hormone, platelet derived growth factor, nerve growth factor, epidermal  
 CC growth factor responsiveness, and for treating diseases associated with  
 CC cell movement. SH2-Bbeta molecule has important implications in cancer  
 CC metastasis, nerve regeneration, angiogenesis and embryogenesis and is  
 CC useful for preventing apoptosis, and treating diabetes, cancer,  
 CC arthritis, immunological diseases, neurological diseases, etc.  
 CC The invention also relates to compositions and methods for identifying  
 CC cytokine, hormone and growth factor signaling pathway agonists and  
 CC antagonists. Human SH2-Bbeta DNA is useful in detecting the presence of  
 CC SH2-Bbeta genes in gene transcripts, for detecting or amplifying DNA's  
 CC with substantial sequence similarity with SH2-Bbeta homologues and  
 CC structural analogues and for gene therapy applications. The present  
 CC sequence is murine JAK2 tyrosine kinase cDNA related to the invention.  
 CC  
 XX SQ Sequence 3495 BP; 1121 A; 688 C; 794 G; 892 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.000114 Length: 3495  
 Score: 16.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x AA085412 (1-3495)

OY 1003 SeraspvaltrpserphcglyvalleuylrgluLeupherThrTyr 1018  
 DB 2695 TCAGATGCTGGAGCTTGGAGTGTCTATACGAACCTTTCACATAC 2742

# RESULT 16

AA085412  
 ID AA085412 standard; cDNA; 3629 BP.

XX AC AA085412;

XX DT 05-OCT-1995 (first entry)

XX DE Murine JAK2 kinase coding sequence and flanking regions.

XX KM JAK family; protein tyrosine kinase; cytokine receptor; mouse;

XX KM phosphorylation; signal transduction; activation; ss.

XX OS Mus musculus.

XX FH Key Location/Qualifiers

XX FT CDS 94..3483

XX FT conflict

XX FT conflict

XX FT conflict

XX FT conflict

XX FT conflict

XX FT conflict

XX FT conflict

XX FT conflict

XX FT conflict

XX FT conflict

XX FT conflict

1089..3579  
 /tag= a  
 /product= Jak2\_kinase  
 /tag= b  
 /note= "Published partial sequence of Jak2 cDNA  
 (Harpur et al., Oncogene 7:1347-1353(1992))  
 has CCC at this position"  
 /tag= c  
 /note= "Published partial sequence of Jak2 cDNA  
 (Harpur et al., Oncogene 7:1347-1353(1992))  
 differs from AA085412 in having the  
 nucleotides shown in brackets at the following  
 positions: 1089(T), 1103(C), 1114(G), 1119(G),  
 1122(C), 1128(C), 1131(G), 1134(G), 1137(C),  
 1140(G), 1143(G), 1146(C), 1188(T), 1194(G),  
 1230(A), 1245(T), 1260(T), 1266(C), 1272(T),  
 1275(A), 1293(T), 1305(T), 1323(C), 1341(A),  
 1344(A), 1359(G), 1365(A), 1368(T), 1374(T),  
 1401(C), 1413(C), 1431(T), 1458(A), 1476(G),  
 1488(T), 1511..1512(GT), 1578(C), 1590(T),

FT 1593(T), 1602(G), 1623(T), 1642(G), 1657(C),  
 FT 1728(G), 1743(C), 1755(C), 1770(A), 1809(G),  
 FT 1816(G), 1821(C), 1857(A), 1878(T), 1935(A),  
 FT 1938(A), 1963(T), 1974(G), 2025(T), 2035(G),  
 FT 2079(C), 2082(C), 2085(C), 2253(A), 2259(G),  
 FT 2283(A), 2285(C), 2433(G), 3453(C), 3579(C),  
 FT 2226  
 /tag= d  
 /note= "location of a 7 amino acid insert detected  
 in previous studies, but not in the present  
 study"  
 FT conflict 3595..3619  
 /tag= e  
 /note= "three extra nucleotides (all A's) were  
 present in previous studies at positions  
 3595, 3598 and 3619 in the 3'-UTR"

PN W09503701-A.

XX 09-FEB-1995.

XX 29-JUL-1994; 94MO-US08676.

XX 29-JUL-1993; 93US-0097997.

XX (SJUD-) SJ JUDE CHILDREN'S RES HOSPITAL.

XX Ihle JN, Quelle FW, Silvennoinen O, Witthuhn BA;

XX WPI; 1995-081950/11.

XX P-PSDB: AAR70830.

XX Inhibiting a cellular response to a cytokine by inhibiting Jak

XX kinase - to treat diseases caused by excessive response to

XX cytokine, e.g. erythrocytosis and other cellular proliferative

XX diseases

XX Claim 29; Fig 1; 167bp; English.

XX Inhibiting the activity of a jak kinase (pref. jak1, jak2, jak3 or

XX tyk2) in a eukaryotic cell is claimed as a method of inhibiting the

XX biological response of that cell to a cytokine (not IL-3 or

XX erythropoietin). The present sequence (murine JAK2 kinase) encodes a

XX amino acid positions 758-776. Antibodies which selectively bind the

XX epitope are able to bind JAK2 without interfering with the activity

XX of the kinase. Such antibodies are claimed and are useful for

XX detecting and extracting JAK2. The 71 nucleotide differences noted

XX between the present sequence and the published partial sequence

XX result in 9 amino acid changes.

XX SQ Sequence 3629 BP; 1142 A; 733 C; 842 G; 912 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 0.000118 Length: 3629

XX Score: 16.00 Matches: 16

XX Percent Similarity: 100.00% Conservative: 0

XX Best Local Similarity: 100.00% Mismatches: 0

XX Query Match: 1.46% Indels: 0

XX DB: 16 Gaps: 0

US-09-397-967-16 (1-1099) x AA085412 (1-3629)

OY 1003 SeraspvaltrpserphcglyvalleuylrgluLeupherThrTyr 1018

DB 3196 TCAGATGCTGGAGCTTGGAGTGTCTATACGAACCTTTCACATAC 3243

# RESULT 17

AA066244  
 ID AA066244 standard; cDNA; 3629 BP.

XX AC AA066244;





```

FT      /*tag- ae
FT      replace (1368, T)
FT      /*tag- af
FT      replace (1374, T)
FT      /*tag- ag
FT      replace (1401, C)
FT      /*tag- ah
FT      replace (1413, C)
FT      /*tag- ai
FT      replace (1431, T)
FT      /*tag- aj
FT      replace (1453, Q)
FT      /*tag- ak
FT      replace (1476, G)
FT      /*tag- al
FT      replace (1488, T)
FT      /*tag- am
FT      replace (1511..1512, GT)
FT      /*tag- an
FT      replace (1578, C)
FT      /*tag- ao
FT      replace (1590, T)
FT      /*tag- ap
FT      replace (1593, T)
FT      /*tag- aq
FT      replace (1602, G)
FT      /*tag- ar
FT      replace (1623, T)
FT      /*tag- as
FT      replace (1642, G)
FT      /*tag- at
FT      replace (1657, C)
FT      /*tag- au
FT      replace (1728, G)
FT      /*tag- av
FT      replace (1743, C)
FT      /*tag- aw
FT      replace (1755, C)
FT      /*tag- ax
FT      replace (1770, A)
FT      /*tag- ay
FT      replace (1809, G)
FT      /*tag- az
FT      replace (1816, G)
FT      /*tag- ba
FT      replace (1821, C)
FT      /*tag- bb
FT      replace (1857, A)
FT      /*tag- bc
FT      replace (1878, T)
FT      /*tag- bd
FT      replace (1935, A)
FT      /*tag- be
FT      replace (1938, A)
FT      /*tag- bf
FT      replace (1963, T)
FT      /*tag- bg
FT      replace (1974, G)
FT      /*tag- bh
FT      replace (2025, T)
FT      /*tag- bi
FT      replace (2055, G)
FT      /*tag- bj
FT      replace (2079, C)
FT      /*tag- bk
FT      replace (2082, C)
FT      /*tag- bl
FT      /*tag- bm
FT      /*note- "Nucleotides corresponding to 7 amino acid insert"
FT      replace (2253, A)
FT      /*tag- bn
FT      replace (2259, G)
FT      conflict

```

```

FT      /*tag- bo
FT      replace (2283, A)
FT      /*tag- bp
FT      replace (2285, C)
FT      /*tag- bq
FT      replace (2433, G)
FT      /*tag- br
FT      replace (3453, C)
FT      /*tag- bs
FT      3'UTR
FT      3484..3629
FT      /*tag- bt
FT      replace (3579, C)
FT      /*tag- bu
FT      replace (3595, A)
FT      /*tag- bv
FT      replace (3599, A)
FT      /*tag- bw
FT      replace (3620, A)
FT      /*tag- bx
FT      conflict

US6210654-B1.
XX      03-APR-2001.
XX      08-OCT-1997; 97US-0946994.
XX      18-JUN-1996; 96US-0665574.
XX      29-JUL-1993; 93US-0097997.
XX      (SUSD-) ST JUDE CHILDREN'S HOSPITAL.
XX      Ihle J, Wittehu BA, Quelle FW, Silvenoinen O;
XX      WPI: 2001-265367/27.
XX      P-PSDB: AAE00352.
XX      Modulating a biological response mediated by Jak kinase 2 activation to
XX      a cytokine, useful for treating excessive proliferation of eukaryotic
XX      cells, comprises inhibiting or enhancing tyrosine kinase activity of
XX      Jak kinase in the cell -.
XX      PS      Example 1; Fig 1; 100pp: English.
XX      CC      The present sequence is a cDNA encoding murine (Janus kinase 2) Jak2
XX      CC      tyrosine kinase. Jak2 sequence has a 600 amino acid long N-terminus that
XX      CC      lacks obvious SH2 (Src homology 2) and SH3 domains. Following this is a
XX      CC      kinase related domain (domain 2) and a carboxyl kinase domain (domain
XX      CC      1). Jak kinases mediate cytokine activity through their tyrosine

Alignment Scores:
Pred. No.: 0.000118      Length: 3629
Score: 16.00      Matches: 16
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.46%      Indels: 0
DB: 22      Gaps: 0

US-09-397-967-16 (1-1099) x AAD03607 (1-3629)
QY      1003 SeraspValTyrSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
DB      3186 TCAGATGTGTGAGAGCTTGGAGTGTCTATACGAACCTTTCACATAC 3243

RESULT 19
AAX80971
ID      AAX80971 standard; cDNA; 4482 BP.
XX      AC      AAX80971;
XX      DT      03-SEP-1999 (first entry)
XX      DE      Human JAK2 kinase encoding cDNA.
XX

```

KW JAK2 kinase; arteriosclerosis; asthma; bronchitis; emphysema; psoriasis;  
 KW inflammatory bowel disease; inflammation; osteoarthritis; oncogenesis;  
 KW rheumatoid arthritis; septic shock; systemic lupus erythematosus;  
 KW Leukemia; human; ss.  
 OS Homo sapiens.  
 XX  
 XX US5914393-A.  
 PN 22-JUN-1999.  
 PD 05-DEC-1995; 95US-0567508.  
 PF 05-DEC-1995; 95US-0567508.  
 PR 05-DEC-1995; 95US-0567508.  
 XX  
 XX (INCY-) INCYTE PHARM INC.  
 PA Coleman R, Stuart SG;  
 PI  
 XX  
 XX WPI; 1999-384188/32.  
 DR P-PSDB; AAY21698.  
 XX  
 PT Polypeptides and amino acids useful for modulating human jak2 kinase  
 PT activity  
 PS  
 XX Claim 2; Columns 23-28; 37pp; English.  
 CC This cDNA encodes a human JAK2 kinase polypeptide. Host cells transformed  
 CC with recombinant jak2 kinase nucleic acid are used for the recombinant  
 CC production of the protein. Purified JAK2 may be used to produce  
 CC antibodies or identify antagonists or inhibitors of JAK2. JAK2, anti-JAK2  
 CC antibodies or JAK2 antagonists or inhibitors may be used to treat,  
 CC prevent or diagnose conditions associated with altered or uncontrolled  
 CC JAK2 expression, e.g. arteriosclerosis, asthma, bronchitis, emphysema,  
 CC inflammatory bowel disease, inflammation, leukemia, oncogenesis,  
 CC osteoarthritis, psoriasis, rheumatoid arthritis, septic shock and  
 CC systemic lupus erythematosus.  
 CC  
 XX  
 SQ Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 0.000144 Length: 4482  
 Score: 16.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: Gaps: 0  
 US-09-397-967-16 (1-1099) x AAX80971 (1-4482)  
 QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTYR 1018  
 Db 3442 TCAGATGTTGGAGCTTGGAGTGTCTGTATGACTTTTCACATAC 3489  
 RESULT 20  
 AAZ58947  
 ID AAZ58947 standard; cDNA; 4482 BP.  
 XX  
 AC AAZ58947;  
 XX  
 DT 03-MAY-2000 (first entry)  
 DE Human JAK2 kinase (HJAK2) encoding cDNA.  
 XX  
 DE Human JAK2 kinase (HJAK2) encoding cDNA.  
 XX  
 DE Janus family nonreceptor protein-tyrosine kinase-2; JAK2; HJAK2; human;  
 KW signal transduction; arteriosclerosis; asthma; bronchitis; emphysema;  
 KW inflammatory bowel disease; leukemia; oncogenesis; osteoarthritis;  
 KW psoriasis; rheumatoid arthritis; systemic lupus erythematosus;  
 KW cytostatic; osteopathic; dermatological; antibacterial; septic shock;  
 KW immunosuppressive; ss.  
 XX  
 OS Homo sapiens.  
 XX

PN US6019966-A.  
 XX  
 XX 01-FEB-2000.  
 PB  
 XX  
 XX 19-NOV-1998; 9805-0196480.  
 PE  
 XX  
 PR 05-DEC-1995; 95US-0567508.  
 XX  
 XX (INCY-) INCYTE PHARM INC.  
 PA Stuart SG, Coleman R;  
 XX  
 XX Stuart SG, Coleman R;  
 PI  
 XX  
 XX WPI; 2000-146859/13.  
 DR P-PSDB; AAY77552.  
 XX  
 XX Human Janus family nonreceptor protein-tyrosine kinases useful as  
 PT diagnostic reagents and for preventing, diagnosing and treating  
 PT diseases such as arteriosclerosis, asthma and leukemia -  
 PS  
 XX Example 5; Fig 1A-F; 33pp; English.  
 CC This cDNA encodes a human Janus family nonreceptor protein-tyrosine  
 CC kinase-2 (JAK2) polypeptide (HJAK2). The JAK2 polypeptides may be used  
 CC as diagnostic reagents as they react with a range of target proteins  
 CC including growth hormone, prolactin, erythropoietin and cytokine  
 CC receptors. They may also be used for the production of antibodies  
 CC specific for JAK2, which may be used to inhibit its activity and prevent  
 CC or treat disorders associated with over expression of JAK2. Conversely,  
 CC the JAK2 polypeptide may be administered to supplement the patients own  
 CC production and counter mutations that may lead to the expression of an  
 CC inactive enzyme. The protein may also be used to screen candidate  
 CC reagents for modulators of JAK2 function. The antagonists and antibodies  
 CC bind to the JAK2 protein and prevent the transfer of high energy  
 CC phosphate molecules, therefore blocking signal transduction. Disorders  
 CC that may be treated by administration of JAK2 polypeptides, anti-JAK2  
 CC antibodies and the agonists and/or antagonists, include arteriosclerosis,  
 CC asthma, bronchitis, emphysema, inflammatory bowel disease, leukemia,  
 CC oncogenesis, osteoarthritis, psoriasis, rheumatoid arthritis, septic  
 CC shock and systemic lupus erythematosus.  
 CC  
 XX  
 SQ Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 0.000144 Length: 4482  
 Score: 16.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.46% Indels: 0  
 DB: Gaps: 0  
 US-09-397-967-16 (1-1099) x AAZ58947 (1-4482)  
 QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTYR 1018  
 Db 3442 TCAGATGTTGGAGCTTGGAGTGTCTGTATGACTTTTCACATAC 3489  
 RESULT 21  
 AAD24311  
 ID AAD24311 standard; DNA; 5117 BP.  
 XX  
 AC AAD24311;  
 XX  
 DT 07-MAR-2002 (first entry)  
 DE Human JAK2 (Janus kinase) DNA.  
 XX  
 DE Human JAK2 (Janus kinase) DNA.  
 XX  
 DE Human; SOCS; suppressor of cytokine signalling; autoimmune disorder; JAK;  
 KW Janus kinase; signal transducer and activator of transcription; Stat;  
 KW rheumatoid arthritis; Wegener's granulomatosis; chronic active hepatitis;  
 KW Crohn's disease; hemolytic anaemia; myasthenia gravis; vasculitis;  
 KW diabetes mellitus; thyroiditis; inflammatory bowel disease; nephrotropic;  
 KW

KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;  
KW hepatotropic; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 478..3876  
FT /\*tag= a  
FT /product= "Human Jak2 protein"  
XX  
XX MO200179555-A2.  
XX  
XX 25-OCT-2001.  
XX  
XX 13-APR-2001; 2001WO-US12131.  
XX  
XX 14-APR-2000; 2000US-0549654.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Hancock MW, Ozkaynak E;  
XX  
XX WPI: 2002-034368/04.  
XX  
XX P-PSDB: AAE15179.  
XX  
XX Monitoring transplant acceptance or autoimmune disease, useful e.g. for  
XX assessing therapy, comprises measuring levels of Stat or their  
XX inhibitors  
XX  
XX Example: Fig 8; 218pp; English.  
XX  
XX The present invention relates to a method for monitoring acceptance of a  
XX transplant or an autoimmune disease in a mammal. The method comprising  
XX determining the amount of at least one of Stat4 (signal transducer and  
XX activator of transcription), Stat6, SOCS1 (suppressor of cytokine  
XX signaling) or SOCS3 mRNA or protein in a sample of the transplant, taken  
XX from the host or an affected tissue sample. Stats are activated by  
XX receptor-associated Janus Kinases (Jaks) which include Jak1, Jak2, Tyk2,  
XX Jak3. The method is used to determine whether acceptance of a transplant  
XX has been induced or to determine if autoimmune disorders (systemic  
XX lupus erythematosus, glomerulonephritis, rheumatoid arthritis, multiple  
XX sclerosis, myasthenia gravis, haemolytic anaemia, nephrotic syndrome,  
XX thyroiditis, diabetes mellitus, Crohn's disease, inflammatory bowel  
XX disease and vasculitis) are being treated successfully and may be used  
XX to adjust treatment regimes. The present sequence is human Jak2 DNA.  
XX  
XX  
XX Sequence 5117 BP; 1623 A; 927 C; 1103 G; 1464 T; 0 other;  
XX  
XX  
XX Alignment Scores:  
XX Pred. No.: 0.000163 Length: 5117  
XX Score: 16.00 Matches: 16  
XX Percent Similarity: 100.00% Conservative: 0  
XX Best Local Similarity: 100.00% Mismatches: 0  
XX Query Match: 1.46% Indels: 0  
XX DB: 24 Gaps: 0  
XX  
XX US-09-397-967-16 (1-1099) x AAD24311 (1-5117)  
XX  
XX QY 1003 SerAapValTrpSerPheGlyValIleuTyrrGluLeuPheThrTyr 1018  
XX ||||||||||||||||||||||||||||||||||||||||||||||||  
XX DB 3580 TCAGATGTTGGAGCTTGACGCTTCGTATGAACCTTCCACATAC 3627  
XX  
XX RESULT 22  
XX ID AAK78096  
XX AAK78096 standard; DNA; 1498 BP.  
XX  
XX AAK78096;  
XX  
XX 07-NOV-2001 (first entry)  
XX  
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32908.  
XX

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
XX Homo sapiens.  
XX  
XX  
XX MO200157182-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX  
XX 17-JAN-2001; 2001WO-US01354.  
XX  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224516.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226868.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0228287.  
XX 01-SEP-2000; 2000US-0228343.  
XX 01-SEP-2000; 2000US-0229344.  
XX 01-SEP-2000; 2000US-0229345.  
XX 05-SEP-2000; 2000US-0229509.  
XX 05-SEP-2000; 2000US-0229513.  
XX 06-SEP-2000; 2000US-0230437.  
XX 06-SEP-2000; 2000US-0230438.  
XX 08-SEP-2000; 2000US-0230438.  
XX 08-SEP-2000; 2000US-0231242.  
XX 08-SEP-2000; 2000US-0231243.  
XX 08-SEP-2000; 2000US-0231244.  
XX 08-SEP-2000; 2000US-0231413.  
XX 08-SEP-2000; 2000US-0231414.  
XX 08-SEP-2000; 2000US-0232080.  
XX 08-SEP-2000; 2000US-0232081.  
XX 12-SEP-2000; 2000US-0232081.  
XX 14-SEP-2000; 2000US-0232397.  
XX 14-SEP-2000; 2000US-0232398.  
XX 14-SEP-2000; 2000US-0232399.  
XX 14-SEP-2000; 2000US-0232400.  
XX 14-SEP-2000; 2000US-0232401.  
XX 14-SEP-2000; 2000US-0233063.  
XX 14-SEP-2000; 2000US-0233064.  
XX 14-SEP-2000; 2000US-0233065.  
XX 21-SEP-2000; 2000US-0234223.  
XX 21-SEP-2000; 2000US-0234274.  
XX

PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.

PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI; 2001-483426/52.  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis  
XX  
PS Disclosure; SEQ ID NO 32908; 3071bp + Sequence Listing; English.  
XX  
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I) by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 1498 BP; 359 A; 412 C; 407 G; 320 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 0.0041 Length: 1498  
Score: 14.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.27% Indels: 0  
DB: Gaps: 0  
  
US-09-397-967-16 (1-1099) x AAK78096 (1-1498)  
  
QY 563 SerpHeLeuGluAlaAlaSerLeuMetSerGlnValSerTyr 576  
|||||  
Db 1206 TCATTCCTGAGGAGCGAGCTGTGATGAGCCAGGTGTGTAC 1247  
  
RESULT 23  
AAQ49750  
ID AAQ49750 standard; DNA; 151 BP.  
XX  
AC AAQ49750;  
XX  
DT 10-MAR-1994 (first entry)  
XX  
DE pTK gene LpTK3 partial sequence.  
XX  
DE pTK; protein tyrosine kinase; catalytic domain; c-kit; NGF;  
KW nerve growth factor receptor; megakaryocyte; lymphocyte;  
KW amplification; primer; polymerase chain reaction; PCR; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9315201-A.  
XX  
PD 05-AUG-1993.  
XX  
PF 22-JAN-1993; 93WO-US00586.  
XX  
PR 22-JAN-1992; 92US-0826935.  
XX  
PA (NMEW-) NEW ENGLAND DEACONESS HOSPITAL.

```

XX  Avraham H, Cowley S, Groopman J, Scadden D;
PI  WPI; 1993-320330/40.
XX
XX  New protein tyrosine kinase genes and proteins encoded by genes -
PI  are of human mega-karyocytic origin
XX
XX  Claim 2; Fig 3B; 60pp; English.
XX
CC  PTK genes were identified using two sets of degenerative
CC  oligonucleotide primers: a first set which amplifies all PTK DNA
CC  segments (AA049743-44), and a second set which amplifies highly
CC  conserved sequences present in the catalytic domain of the c-kit
CC  subgroup of PTKs (AA049745-46). The PTK genes identified are described
CC  in AA049747-57 and AAR41897-02.
CC  The LPTKs are expressed in lymphocytic cells, as well as
CC  megakaryocytic cells. The LPTK3 expression prod. exhibited
CC  significant sequence homology with known protein tyrosine kinases
CC  of the NGF receptor family. The protein sequence corresp. to AA049750
CC  is claimed (claim 7) and stated as given in the specification, however
CC  is missing from the publication.
XX
SQ  Sequence 151 BP; 37 A; 44 C; 41 G; 29 T; 0 other;

Alignment Scores:
Pred. No.:      0.0379      Length:      151
Score:          12.00      Matches:      12
Percent Similarity: 100.00%  Conservat: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    1.09%      Indels:      0
DB:             14         Gaps:        0

US-09-397-967-16 (1-1099) x AA049750 (1-151)

OY  941 ValHisArgAspLeuAlaIArgAsnIleLeuVal 952
DB  1 GTGCACAGGAGATCTCGGCTCGAACAATCTCTGTC 36

RESULT 24
AAT03093
ID  AAT03093 standard; DNA; 151 BP.
XX
AC  AAT03093;
XX
DT  14-FEB-1996 (first entry)
XX
DE  Protein tyrosine-kinase LPTK3 DNA fragment.
XX
KW  Protein tyrosine-kinase; PTK; LPTK3; agonist; cell growth;
KW  differentiation; ss.
XX
OS  Homo sapiens.
XX
FH  Key      Location/Qualifiers
FT  CDS      1..150
FT           /*tag= a
XX
XX  MO9527061-A1.
XX
XX  12-OCT-1995.
XX
XX  04-APR-1995; 95WO-US04228.
XX
XX  04-APR-1994; 94US-0222616.
XX
XX  (GENE ) GENENTECH INC.
XX
PI  Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
PI  Wood WI;
XX
XX  WPI; 1995-366160/47.
XX
XX  P-PSDB; AAR85932.

```

```

XX  Agonist antibodies which activate specific protein tyrosine
PI  kinase(s) - also activate chimeric proteins of kinase extracellular
PI  domain and Ig constant domain, useful for studying, and therapeutic
PI  modulation of, cell growth and differentiation
XX
XX  Disclosure; Page 37; 125pp; English.
XX
XX  DNA probes based on protein tyrosine-kinase (PTK) sequences were used
XX  to screen cDNA libraries to identify novel PTK genes. A LPTK3 gene
XX  fragment (AAT03093) was isolated from lymphocytic and megakaryocytic
XX  cell line libraries and encoded a peptide (AAR85932) showing homology
XX  to known PTKs. The gene fragment can be used to identify other
XX  new PTK genes, or to design drugs, peptides or antisense constructs
XX  that modulate PTK activity.

SQ  Sequence 151 BP; 37 A; 44 C; 41 G; 29 T; 0 other;

Alignment Scores:
Pred. No.:      0.0379      Length:      151
Score:          12.00      Matches:      12
Percent Similarity: 100.00%  Conservat: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    1.09%      Indels:      0
DB:             16         Gaps:        0

US-09-397-967-16 (1-1099) x AAT03093 (1-151)

OY  941 ValHisArgAspLeuAlaIArgAsnIleLeuVal 952
DB  1 GTGCACAGGAGATCTCGGCTCGAACAATCTCTGTC 36

RESULT 25
AAS10805
ID  AAS10805 standard; cDNA; 214 BP.
XX
AC  AAS10805;
XX
DT  24-OCT-2001 (first entry)
XX
DE  Human Janus kinase 3 (JAK3) RT-PCR fragment #2.
XX
KW  Human; Janus kinase 3; JAK3; RT-PCR; ss; JAK/STAT inhibitor;
KW  signal transducer and activator of transcription; osteoarthritis;
KW  degenerative joint disease; rheumatoid arthritis; leprosy; asthma;
KW  cancer; tumour; leukaemia.
XX
OS  Homo sapiens.
XX
XX  WO200152892-A2.
XX
XX  26-JUL-2001.
XX
XX  22-JAN-2001; 2001WO-US02033.
XX
XX  24-JAN-2000; 2000US-0177872.
XX
XX  28-NOV-2000; 2000US-0723490.
XX
XX  (GENE ) GENZYME CORP.
XX
XX  Vasios G;
XX
XX  WPI; 2001-465338/50.
XX
XX  Use of inhibitors of Janus kinase/signal transducers and activators of
XX  transcription for inhibiting onset and progression of degenerative
XX  joint diseases or disorders such as osteoarthritis, rheumatoid
XX  arthritis -
XX
XX  Example 3; Fig 2; 55pp; English.
XX
XX  The sequence is obtained from an RT (reverse transcriptase)-PCR fragment
XX  from human Janus kinase 3 (JAK3). The invention relates to the use of

```

CC JAK/STAT (Janus kinase/signal transducer and activator of  
CC transcription) inhibitors other than dephosphorylated kinase (DBH) and  
CC hemealdehyde (H) for inhibiting the progression or the likelihood of  
CC developing diseases involving cartilage degradation, and for regulating  
CC the expression of pro-inflammatory agents or cytokines in a chondrocyte,  
CC and cartilage degrading enzymes in a cell. A JAK3/STAT inhibitor of  
CC invention is useful for inhibiting progression or likelihood of  
CC developing osteoarthritis or rheumatoid arthritis. The inhibitor is also  
CC useful for treating other JAK/STAT-mediated diseases or disorders,  
CC including T cell-mediated disorders, mast cell-mediated disorders,  
CC type 2 (cytokine hypersensitivity) disorders, B cell lymphoma, and  
CC myeloid diseases. T cell-mediated disorders include human T cell  
CC leukemia/lymphoma virus (HTLV)-1, Sezary's syndrome, c-abl  
CC transformation, natural killer-like T cell lymphomas (NK-like tumours)  
CC and graft-vs-host disease; cytokine hypersensitivity disorders include  
CC leishmaniasis, leprosy, allergy and viral infections; mast cell-mediated  
CC disorders include hay fever, asthma, hives and anaphylaxis; and  
CC leukemias and lymphomas include acute lymphocytic and lymphoblastic  
CC leukemias, B cell lymphomas and leukemias of myeloid origin. DBH and H  
CC are useful as therapeutic agents in cancers in which JAK3 plays a role  
CC in the initiation or progression of tumorigenesis.

US-09-397-967-16 (1-1099) x AAS10805 (1-214)

Alignment Scores:  
Pred. No.: 0.0526 Length: 214  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x AAS10805 (1-214)  
QY 938 ArgArGcysValHisArgSpLeuAlaAlaArgasn 949  
DB 128 CGCCGCTCGTGACCGCGACCTGCGCCGCCGAAC 163

RESULT 26

ABA50838  
ID ABA50838 standard; DNA; 294 BP.

AC ABA50838;

DT 01-FEB-2002 (first entry)

DE Human breast cell single exon nucleic acid probe #9533.

KW Human; microarray; single exon probe; gene expression; breast;  
KM disease; cancer; ss.

OS Homo sapiens.

PN WO200157271-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00662.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX PA  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2001-496933/54.  
XX

PT New spatially-addressable set of single exon nucleic acid probes,  
PT useful for measuring gene expression in sample derived from human  
PT breast, comprises number of single exon nucleic acid probes  
XX  
XX  
PS Claim 4; SEQ ID NO 9533; 327pp + sequence listing; English.

CC The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BT 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a single exon nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

US-09-397-967-16 (1-1099) x ABA50838 (1-294)

Alignment Scores:  
Pred. No.: 0.0708 Length: 294  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x ABA50838 (1-294)

QY 941 ValHisArgSpLeuAlaAlaArgasn1leuVal 952

DB 91 GTGCATCGTGAATGCGCGACGGAACATCCGTG 126

RESULT 27

ABA68808  
ID ABA68808 standard; DNA; 294 BP.

AC ABA68808;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #17113.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX PA  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2001-496933/54.  
XX

DR WPI: 2001-483447/52.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human fetal liver -  
 XX  
 PS Claim 4: SEQ ID NO 17113; 639pp + sequence listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC fetal liver. The present sequence is a single exon nucleic acid  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.0708 Length: 294  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-397-967-16 (1-1099) x ABA68808 (1-294)  
 QY 941 ValHsArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 Db 91 GTGCATCGTGAATGCGCCGACGACATCTCTG 126  
 RESULT 28  
 ABA35766  
 ID ABA35766 standard; DNA: 294 BP.  
 XX  
 AC ABA35766;  
 XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Probe #14232 for gene expression analysis in human heart cell sample.  
 XX  
 KW Human: gene expression; heart; microarray; vascular system; probe;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157274-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00666.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI: 2001-488899/53.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT hearts -  
 XX  
 PS Claim 4; SEQ ID NO 14232; 530pp; English.

XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart. The  
 CC present sequence is one such probe. The probes may be used for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from the human heart via microarrays. By measuring gene expression, the  
 CC probes are useful for predicting, diagnosing, grading, staging, the  
 CC monitoring and prognosing diseases of the human heart and vascular system  
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
 CC congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.0708 Length: 294  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-397-967-16 (1-1099) x ABA35766 (1-294)  
 QY 941 ValHsArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 Db 91 GTGCATCGTGAATGCGCCGACGACATCTCTG 126  
 RESULT 29  
 AAK17149  
 ID AAK17149 standard; DNA: 294 BP.  
 XX  
 AC AAK17149;  
 XX  
 DT 05-NOV-2001 (first entry)  
 XX  
 DE Human brain expressed single exon probe SEQ ID NO: 17140.  
 XX  
 KW Human: brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI: 2001-483446/52.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains -  
 XX  
 PS Example 4; SEQ ID NO: 17140; 650pp + Sequence Listing; English.  
 CC  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is one of the probes of the  
CC invention.

SO Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;

Alignment Scores:

Alignment Scores:	0.0708	Length:	294
Pred. No.:	12.00	Matches:	12
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	1.09%	Indels:	0
Query Match:	22	Gaps:	0

US-09-397-967-16 (1-1099) x AAK17149 (1-294)

OY 941 VALHISARGSPLEUALAALARGASNI1leuVal 952  
|||||  
DB 91 GTGCATCGTGAATCTGCGCCGACGGAACATCCTGCTG 126

RESULT 30

AAK42934  
ID AAK42934 standard; DNA; 294 BP.

AC AAK42934;

XX 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 17491.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

OS Homo sapiens.

XX WO200157276-A2.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001MO-US00668.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488900/53.

DR WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow.

XX Example 4; SEQ ID NO: 17491; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in bone marrow

CC bone marrow. They can be used to measure gene expression and treatment of cancers

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukemia and myeloma. The present sequence is one of

CC the probes of the invention.

Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	22	Gaps:	0

US-09-397-967-16 (1-1099) x AAK42934 (1-294)

OY 941 VALHISARGSPLEUALAALARGASNI1leuVal 952  
|||||

DB 91 GTGCATCGTGAATCTGCGCCGACGGAACATCCTGCTG 126

RESULT 31

AAI23695  
ID AAI23695 standard; DNA; 294 BP.

AC AAI23695;

XX 12-OCT-2001 (first entry)

DE Probe #13628 for gene expression analysis in human cervical cell sample.

XX Probe; human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer; ss.

OS Homo sapiens.

XX WO200157278-A2.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001MO-US00670.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488901/53.

DR WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells -

XX Claim 25; SEQ ID No 13628; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes

CC (SEMP). The present sequence is one such probe. The SEMPs are derived

CC from human HeLa cells. The SEMPs can be used to produce a single exon

CC microarray, which can be used for measuring human gene expression in a

CC sample derived from human cervical epithelial cells. By measuring gene

CC expression, the probes are therefore useful in grading and/or staging

CC of diseases of the cervix, notably cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pcl\_sequences.

XX Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;

Alignment Scores: 0.0708 Length: 294

Pred. No.: 12.00 Matches: 12

Score: 100.00% Conservative: 0

Percent Similarity: 100.00% Mismatches: 0

Best Local Similarity: 1.09% Indels: 0

Query Match: 22 Gaps: 0



US-09-397-967-16 (1-1099) x AA123695 (1-294)

OY 941 ValHisArgAspLeuAlaIaArgAsnIleLeuVal 952  
 |||||||  
 DB 91 GTGCATCGTATCTGCCCGACGACATCTCTGTTG 126

RESULT 32  
 AA149009  
 ID AA149009 standard; DNA; 294 BP.  
 AC AA149009;  
 XX  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Probe #17695 used to measure gene expression in human placenta sample.  
 XX  
 KW Probe; microarray; human; placenta; antenatal diagnosis;  
 KM genetic disorder; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488897/53.  
 XX  
 DR Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 XX  
 PS Claim 25; SEQ ID No 17695; 654bp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SENP).  
 CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders.  
 CC  
 XX  
 SQ Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;  
 XX

Alignment Scores:  
 Pred. No.: 0.0708 Length: 294  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x AA149009 (1-294)

OY 941 ValHisArgAspLeuAlaIaArgAsnIleLeuVal 952  
 |||||||  
 DB 91 GTGCATCGTATCTGCCCGACGACATCTCTGTTG 126

RESULT 33  
 AA109314  
 ID AA109314 standard; DNA; 294 BP.  
 AC AA109314;  
 XX

XX  
 DT 09-OCT-2001 (first entry)  
 XX  
 DE Probe #9305 used to measure gene expression in human breast sample.  
 XX  
 KW Probe; human; breast disease; breast cancer; development disorder; ss;  
 KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157270-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 29-JAN-2001; 2001WO-US00661.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-476286/51.  
 XX  
 DR Novel single exon nucleic acid probe used to measuring gene expression  
 PT in a human breast -  
 XX  
 PS Claim 25; SEQ ID No 9305; 322bp; English.  
 XX  
 CC The present invention relates to novel single exon nucleic acid probes.  
 CC The present sequence is one such probe. The probes are useful for  
 CC measuring human gene expression in a human breast sample, where the probe  
 CC hybridises at high stringency to a nucleic acid expressed in the human  
 CC breast. The probes are useful for predicting, diagnosing, grading,  
 CC staging, monitoring and prognosing diseases of the human breast,  
 CC particularly those diseases with polygenic aetiology. The diseases  
 CC include: breast cancer; disorders of development, inflammatory diseases  
 CC of the breast; fibrocystic changes; proliferative breast disease and  
 CC non-carcinoma tumours.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published-pct-sequences.  
 CC  
 XX  
 SQ Sequence 294 BP; 69 A; 60 C; 79 G; 86 T; 0 other;  
 XX

Alignment Scores:  
 Pred. No.: 0.0708 Length: 294  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x AA109314 (1-294)

OY 941 ValHisArgAspLeuAlaIaArgAsnIleLeuVal 952  
 |||||||  
 DB 91 GTGCATCGTATCTGCCCGACGACATCTCTGTTG 126

RESULT 34  
 ABS16994  
 ID ABS16994 standard; DNA; 294 BP.  
 AC ABS16994;  
 XX  
 DT 19-AUG-2002 (first entry)  
 XX

DE Human genome-derived single exon probe ORF from lung SEQ ID NO 16985.  
 KW Human: ds; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagazer syndrome;  
 KW pulmonary alveolar proteinosis; fibrocytic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease; open reading frame; ORF.  
 XX Homo sapiens.  
 OS  
 XX MO200186003-A2.  
 PN  
 XX 15-NOV-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US00665.  
 PF  
 XX 04-FEB-2000; 2000US-180312P.  
 PR 26-MAY-2000; 2000US-207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2002-114183/15.  
 DR  
 XX Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples -  
 XX  
 XX  
 PS Claim 4; SEQ ID NO 16985; 634pp; English.  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of  
 CC probes. The novel set of probes which hybridize at high stringency to a  
 CC nucleic acid expressed in the human lung; measuring gene expression in a  
 CC sample derived from human lung, comprising (a) contacting the array with  
 CC a collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of  
 CC the array; identifying exons in a eukaryotic genome, comprising  
 CC (a) algorithmically predicting at least one exon from genomic sequences  
 CC of the eukaryote; and (b) detecting specific hybridization of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene  
 CC expression analysis, and for identifying exons in a gene, particularly  
 CC using human lung derived mRNA and for the study of lung diseases  
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Karagazer syndrome, fibrocytic  
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension

CC and hyaline membrane disease. The present sequence is a single exon  
 CC probe open reading frame of the invention.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 294 BP: 69 A; 60 C; 79 G; 86 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.0708 Length: 294  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-397-967-16 (1-1099) x ABS16994 (1-294)  
 QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 |||||||||||||||||||||||||||||||||||  
 Db 91 GTGCATCTGATCTGCGCCGACGACATCTCTGTG 126  
 RESULT 35  
 ABA45718  
 ID ABA45718 standard; DNA; 367 BP.  
 XX  
 AC ABA45718;  
 XX  
 DT 01-FEB-2002 (first entry)  
 DE  
 XX Human breast cell single exon nucleic acid probe #4413.  
 XX  
 KW Human; microarray; single exon probe; gene expression; breast;  
 KW disease; cancer; ss.  
 XX  
 OS Homo sapiens.  
 OS  
 XX  
 XX WO200157271-A2.  
 PN  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00662.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2001-496933/54.  
 DR  
 XX  
 PT New spatially-addressable set of single exon nucleic acid probes,  
 PT useful for measuring gene expression in sample derived from human  
 PT breast, comprises number of single exon nucleic acid probes -  
 XX  
 PS Claim 1; SEQ ID NO 4413; 327pp + sequence listing; English.  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 474 cells. The method involves contacting  
 CC the probes with a collection of detectably labeled nucleic acids  
 CC derived from mRNA of human breast, and then measuring the label  
 CC bound to each probe of the microarray. The probes are useful for  
 CC verifying the expression of regions of genomic DNA predicted to  
 CC encode proteins. They are useful for gene discovery, and for  
 CC determining predisposition and/or prognosing breast disease. Gene

CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a single exon nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;

Alignment Scores:

Pred. No.:	0.0872	Length:	367
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	22	Gaps:	0

US-09-397-967-16 (1-1099) x ABA5718 (1-367)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952  
Db 289 GTGCATCGTGTGCGCCGACGACATCTGCTG 324

RESULT 36  
ABA56224

ID ABA56224 standard; DNA; 367 BP.

XX ABA56224;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #4529.

XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human foetal liver -

PS Claim 1; SEQ ID NO 4529; 639bp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for  
XX measuring human gene expression in a sample derived from human foetal  
XX liver. The single exon nucleic acid probes may be used for predicting,  
XX measuring and displaying gene expression in samples derived from human  
XX foetal liver. The present sequence is a single exon nucleic acid  
XX probe of the invention.

CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;

Alignment Scores:

Pred. No.:	0.0872	Length:	367
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	22	Gaps:	0

US-09-397-967-16 (1-1099) x ABA56224 (1-367)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952  
Db 289 GTGCATCGTGTGCGCCGACGACATCTGCTG 324

RESULT 37  
ABA25869

ID ABA25869 standard; DNA; 367 BP.

XX ABA25869;

DT 23-JAN-2002 (first entry)

DE Probe #4335 for gene expression analysis in human heart cell sample.

XX Human; gene expression; heart; microarray; vascular system; probe;

XX cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease; ss.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human  
XX hearts -

PS Claim 1; SEQ ID NO 4335; 530bp; English.

XX The present invention relates to single exon nucleic acid probes for  
XX measuring human gene expression in a sample derived from human heart. The  
XX present sequence is one such probe. The probes may be used for  
XX predicting, measuring and displaying gene expression in samples derived  
XX from the human heart via microarrays. By measuring gene expression, the  
XX probes are useful for predicting, diagnosing, grading, staging,  
XX monitoring and prognosing diseases of the human heart and vascular system  
XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
XX congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;

Alignment Scores:

Pred. No.:	0.0872	Length:	367
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	22	Gaps:	0

US-09-397-967-16 (1-1099) x ABA25869 (1-367)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 |||||  
 Db 289 GTGCATCGTGCATCTGGCCGACGAAACATCTCGTGTG 324

RESULT 38  
 AAK04412  
 ID AAK04412 standard; DNA; 367 BP.

AC AAK04412;  
 DT 05-NOV-2001 (first entry)  
 DE Human brain expressed single exon probe SEQ ID NO: 4403.  
 XX  
 XX Human; brain expressed exon; gene expression analysis; probe;  
 KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KM epilepsy; cancer; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-483446/52.  
 DR  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains -  
 PS  
 XX Example 4; SEQ ID NO: 4403; 650pp + Sequence Listing; English.  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is one of the probes of the  
 CC invention.  
 CC  
 XX  
 XX Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;  
 SQ

Alignment Scores:

Pred. No.:	0.0872	Length:	367
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	22	Gaps:	0

US-09-397-967-16 (1-1099) x AAK04412 (1-367)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 |||||  
 Db 289 GTGCATCGTGCATCTGGCCGACGAAACATCTCGTGTG 324

RESULT 39  
 AAK29908  
 ID AAK29908 standard; DNA; 367 BP.

AC AAK29908;  
 DT 06-NOV-2001 (first entry)  
 DE Human bone marrow expressed single exon probe SEQ ID NO: 4465.  
 XX  
 XX Human bone marrow expressed single exon probe SEQ ID NO: 4465.  
 DE Human; bone marrow expressed exon; gene expression analysis; probe;  
 KM microarray; cancer; leukemia; lymphoma; myeloma; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 XX WPI; 2001-488900/53.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human bone marrow -  
 PS  
 XX Example 4; SEQ ID NO: 4465; 658pp + Sequence Listing; English.  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of  
 CC the probes of the invention.  
 CC  
 XX  
 XX Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;  
 SQ

Alignment Scores:

Pred. No.:	0.0872	Length:	367
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	22	Gaps:	0

US-09-397-967-16 (1-1099) x AAK29908 (1-367)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 |||||  
 Db 289 GTGCATCGTGCATCTGGCCGACGAAACATCTCGTGTG 324

RESULT 40  
 AAI14495  
 ID AAI14495 standard; DNA; 367 BP.

```
XX AC AAI14495;
XX AC
XX DT 12-OCT-2001 (first entry)
XX DE Probe #4428 for gene expression analysis in human cervical cell sample.
XX KW Probe: human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer; ss.
XX OS Homo sapiens.
XX PN WO200157278-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00670.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI: 2001-488901/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human cervical epithelial cells -
XX PS Claim 25; SEQ ID NO 4428; 487bp; English.
XX SC The present invention relates to human single exon nucleic acid probes
XX CC (SNP). The present sequence is one such probe. The SNPs are derived
XX CC from human HeLa cells. The SNPs can be used to produce a single exon
XX CC microarray, which can be used for measuring human gene expression in a
XX CC sample derived from human cervical epithelial cells. By measuring gene
XX CC expression, the probes are therefore useful in grading and/or staging
XX CC of diseases of the cervix, notably cervical cancer.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SC Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;
XX SC
XX SC Alignment Scores:
XX SC Pred. No.: 0.0872 Length: 367
XX SC Score: 12.00 Matches: 12
XX SC Percent Similarity: 100.00% Conservative: 0
XX SC Best Local Similarity: 100.00% Mismatches: 0
XX SC Query Match: 1.09% Indels: 0
XX SC DB: 22 Gaps: 0
XX SC
XX SC US-09-397-967-16 (1-1099) x AAI14495 (1-367)
OY 941 ValHISArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 289 GTGCATCTGATCTGGCCGACGGAACATCTCTGGTG 324
RESULT 41
AAI35871
ID AAI35871 standard; DNA; 367 BP.
XX AC AAI35871;
XX AC
XX DT 17-OCT-2001 (first entry)
XX DT
XX PF Probe #4557 used to measure gene expression in human placenta sample.
```

```
XX AC
XX AC Probe: microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder; ss.
XX KW
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00663.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI: 2001-488997/53.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human placenta -
XX PS Claim 25; SEQ ID NO 4557; 654bp; English.
XX SC The present invention relates to single exon nucleic acid probes (SNP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders.
XX SC Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;
XX SC
XX SC Alignment Scores:
XX SC Pred. No.: 0.0872 Length: 367
XX SC Score: 12.00 Matches: 12
XX SC Percent Similarity: 100.00% Conservative: 0
XX SC Best Local Similarity: 100.00% Mismatches: 0
XX SC Query Match: 1.09% Indels: 0
XX SC DB: 22 Gaps: 0
XX SC
XX SC US-09-397-967-16 (1-1099) x AAI35871 (1-367)
OY 941 ValHISArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 289 GTGCATCTGATCTGGCCGACGGAACATCTCTGGTG 324
RESULT 42
AAI04322
ID AAI04322 standard; DNA; 367 BP.
XX AC AAI04322;
XX AC
XX DT 09-OCT-2001 (first entry)
XX DT
XX PF Probe #4313 used to measure gene expression in human breast sample.
XX KW Probe: human; breast disease; breast cancer; development disorder; ss;
XX KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX OS Homo sapiens.
XX PN WO200157270-A2.
XX PD 09-AUG-2001.
XX PF 29-JAN-2001; 2001WO-US00661.
```

XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2001-476286/51.  
 XX  
 PT Novel single exon nucleic acid probe used to measuring gene expression  
 PT in a human breast -  
 XX  
 PS Claim 25; SEQ ID NO 4313; 322pp; English.  
 XX  
 CC The present invention relates to novel single exon nucleic acid probes.  
 CC The present sequence is one such probe. The probes are useful for  
 CC measuring human gene expression in a human breast sample, where the probe  
 CC hybridizes at high stringency to a nucleic acid expressed in the human  
 CC breast. The probes are useful for predicting, diagnosing, grading,  
 CC staging, monitoring and prognosing diseases of the human breast,  
 CC particularly those diseases with polygenic aetiology. The diseases  
 CC include: breast cancer, disorders of development, inflammatory diseases  
 CC of the breast, fibrocystic changes, proliferative breast disease and  
 CC non-carcinoma tumours.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).  
 XX  
 SQ Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0.0872 Length: 367  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: Gaps: 0  
 US-09-397-967-16 (1-1099) x AA104322 (1-367)  
 QY 941 VAlHtSArGASpLeAlaAlaArGASnLleuVal 952  
 DB 289 GTGCATCGTGTGGCCGACGAGACATCTGTG 324  
 XX  
 RESULT 43  
 ABS04474  
 ID ABS04474 standard; DNA; 367 BP.  
 XX  
 AC ABS04474:  
 XX  
 DT 19-AUG-2002 (first entry)  
 XX  
 DE Human genome-derived single exon probe from lung SEQ ID No 4465.  
 XX  
 KW Human: ds; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 PN WO200186003-A2.

XX 15-NOV-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US00665.  
 PF  
 XX  
 XX 04-FEB-2000; 2000US-180312P.  
 PR 26-MAY-2000; 2000US-207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2002-114183/15.  
 XX  
 PT Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples -  
 XX  
 PS Claim 1; SEQ ID NO 4465; 634pp; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of  
 CC probes; the novel set of probes which hybridize at high stringency to a  
 CC nucleic acid expressed in the human lung; measuring gene expression in a  
 CC sample derived from human lung, comprising (a) contacting the array with  
 CC a collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of  
 CC the array; identifying exons in a eukaryotic genome, comprising  
 CC (a) algorithmically predicting at least one exon from genomic sequences  
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene  
 CC expression analysis, and for identifying exons in a gene, particularly  
 CC using human lung derived mRNA and for the study of lung diseases  
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic  
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
 CC and hyaline membrane disease. The present sequence is a single exon  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).  
 XX  
 SQ Sequence 367 BP; 102 A; 71 C; 89 G; 105 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0.0872 Length: 367  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0

DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x ABS04474 (1-367)

OY 941 VALHISARGSPLEUALAALARGASNIILEUVAL 952  
 |||||||||||||||||||||||||||||||||||

DB 289 GTGCATCGGACTGTGCGCAGCAGAACATCTCGTGA 324

RESULT 44  
 AAA98758  
 ID AAA98758 standard; DNA; 405 BP.

AC AAA98758;

DT 07-FEB-2001 (first entry)

DE Human RET proto-oncogene exon 15 DNA.

KM RET proto-oncogene; human; cytostatic; obstipation;  
 KM multiple endocrine neoplasia syndrome type 2A; undine syndrome;  
 KM familial medullary thyroid gland carcinoma; sudden infant death;  
 KM central breathing regulation disorder; ds.

OS Homo sapiens.

PN DEJ9910912-A1.

PD 21-SEP-2000.

PF 11-MAR-1999; 99DE-1010912.

PR 11-MAR-1999; 99DE-1010912.

PA (UYDR ) UNIV DRESDEN TECH.

PI Fltze G, Schackert HK, Roegner D;

DR WPI; 2000-588405/56.

PT New human RET proto-oncogene variants for determining disease  
 PT disposition and tailoring specific individual therapies, for e.g.  
 PT multiple endocrine neoplasia syndrome type 2A or for familial medullary  
 PT thyroid gland carcinoma

PS Disclosure: Flg 21; 14pp; German.

CC This invention describes novel human RET proto-oncogene variants  
 CC which have cytoskeletal activity. The proto-oncogenes are used to  
 CC identify dispositions to forms of disturbance or idiopathic obstipation,  
 CC to determine disposition for multiple endocrine neoplasia syndrome type  
 CC 2A, familial medullary thyroid gland carcinoma, central breathing  
 CC regulation disorder, in particular undine syndrome or sudden infant  
 CC death. They can be used to characterize and detect homozygous variants  
 CC for position 135A of RET proto-oncogene, optionally with other genetic  
 CC characteristics. Heterozygous variants, e.g. containing a variation in  
 CC the cysteine rich region of RET (1.e. position 1825) and at position  
 CC 135A, can be identified. The sequence variants can be used for  
 CC development of therapeutics, especially new classes of therapeutics,  
 CC targeted to the human RET proto-oncogene, its 5' regulatory region or  
 CC promoter and regulators of transcription and translation. They can be  
 CC used to individually optimize therapy or intervention targeted to the  
 CC Ret receptor tyrosine kinase. The sequences can be used to construct  
 CC vectors, in particular to develop pharmaceutically relevant agents and  
 CC for diagnostic kits, especially for genotyping. The variants can also  
 CC be used to develop in vitro, preferably in cell culture, and in vivo,  
 CC transgenic animals and test systems, for expression of individual forms  
 CC of the human RET proto-oncogene, where the test system is used to look  
 CC at pathophysiology of disease and general medical characteristics  
 CC associated with RET proto-oncogene and to develop and test individually  
 CC specific therapeutics.

Sequence 405 BP; 85 A; 128 C; 117 G; 75 T; 0 other;

# Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
12.00	100.00%	405	12	0	0	0	0
Best Local Similarity:	100.00%						
Query Match:	1.09%						
DB:	21						

US-09-397-967-16 (1-1099) x AAA98758 (1-405)

OY 941 VALHISARGSPLEUALAALARGASNIILEUVAL 952  
 |||||||||||||||||||||||||||||||||||

DB 175 GTTCATCGGACTGTGCGCAGCAGAACATCTCGTGA 210

RESULT 45

AAC95082  
 ID AAC95082 standard; cDNA; 917 BP.

AC AAC95082;

DT 19-FEB-2001 (first entry)

DE Cat flea hindgut and Malpighian tubule (HMT) cDNA, SEQ ID NO:1577.

KM Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;

KM flea infestation; vaccine; antiparasitic; therapeutic target;

KM diagnosis; detection; ss.

OS Ctenocephalides felis.

PN WO200061621-A2.

PD 19-OCT-2000.

PF 07-APR-2000; 2000WO-US09437.

PR 09-APR-1999; 99US-0128704.

PA (HESK-) HESKA CORP.

PI Brandt KS, Galnes PJ, Stinchcomb DT, Wisniewski N;

DR WPI; 2000-656323/63.

PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic  
 PT acids useful for the prevention, diagnosis and treatment of flea  
 PT infestations -

PS Claim 26; Page 768; 964pp; English.

CC The invention relates to novel cat flea (Ctenocephalides felis) nucleic  
 CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue  
 CC or head and nerve cord (HNC) tissue. The invention also relates to the  
 CC encoded proteins. The invention additionally encompasses expression  
 CC constructs, recombinant viruses and recombinant cells comprising the  
 CC nucleic acids of the invention, recombinant production of the proteins,  
 CC antibodies against the proteins, a method of identifying inhibitors of  
 CC the proteins, and compositions comprising the inhibitors for  
 CC administration to an animal. The nucleic acids, and the proteins they  
 CC encode may be used in the prevention, treatment and diagnosis of diseases  
 CC associated with flea infestations. For example, the nucleic acids may be  
 CC used to produce an HMT or HNC protein according to standard recombinant  
 CC DNA methodology by inserting the nucleic acids into a host cell and  
 CC culturing the cell to express the protein. The HMT and HNC nucleic acids  
 CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect  
 CC and quantitate the presence of cat flea or other homologous nucleic acid  
 CC sequences in samples. They may also be used to study the expression and  
 CC function of the proteins and their role in metabolism. The HMT and HNC  
 CC proteins may be used as antigens in the production of specific  
 CC antibodies, and in assays to identify modulators (agonists and  
 CC antagonists) of HMT and/or HNC protein expression and activity. The  
 CC anti-HMT/HNC protein antibodies and antagonists may also be used to  
 CC downregulate protein expression and activity. The antibodies may also be

CC used as diagnostic agents for detecting the presence of flea polypeptides  
CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The  
CC present sequence represents a cat flea HMT cDNA of the invention.

SO Sequence 917 BP; 263 A; 175 C; 207 G; 242 T; 30 other;

Alignment Scores:

Pred. No.:	0.206	Length:	917
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	21	Gaps:	0

US-09-397-967-16 (1-1099) x AAC95082 (1-917)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 169 GTCCATCGAGATTGGCTGCCAGAAATATCTGTG 204

RESULT 46

ID AAS87117 standard; cDNA; 1117 BP.

XX AAS87117:

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #22921.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX MO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR P-PSDB; ABG22930.

PT New isolated polynucleotide and encoded polypeptides; useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

PS Claim 1; SEQ ID No 22921; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful for treating

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human

CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pat\_sequences.

SO Sequence 1117 BP; 308 A; 246 C; 292 G; 271 T; 0 other;

Alignment Scores:

Pred. No.:	0.248	Length:	1117
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	23	Gaps:	0

US-09-397-967-16 (1-1099) x AAS87117 (1-1117)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 235 GTCCATCGATGATCTGCCAGCAACATCTGTG 270

RESULT 47

ID AAO53470 standard; DNA; 1509 BP.

XX AAO53470:

DT 16-JUN-1994 (first entry)

DE PKR gene B1.

XX Lambda gt11; expression vector; lambda-B1-Elk; protein tyrosine kinase;

KW Elk; B1; Eph; subfamily; receptor-like tyrosine kinase; eph; eck;

KW phosphorylation; phosphorylated kinase insert domain; growth factor;

KW receptor kinase; platelet-derived growth factor receptor; ss.

XX Rattus rattus.

XX CA2083521-A.

XX 01-OCT-1993.

XX 23-NOV-1992; 92CA-2083521.

XX 31-MAR-1992; 92US-0861390.

PA (MOUN ) MOUNT SINAI HOSPITAL CORP.

PI Letwin K, Pawson A, Reedijk M;

DR WPI; 1993-406300/51.

DR P-PSDB; AAR44512.

PT Expression of phosphorylated exogenous protein - in host cells

PT transformed with two vectors, one for the protein, the other for

PT catalytic domain of protein kinase

PS Disclosure; Fig 1; 55pp; English.

XX This sequence represents a fragment of the lambda gt11 expression

XX vector, lambda-B1-Elk, which encodes the catalytic sequence of the

XX protein tyrosine kinase, Elk. The Elk gene, B1, encodes a protein

XX which is a member of the Eph subfamily of protein tyrosine kinases.

XX The Elk product is very similar to two other receptor-like tyrosine

XX kinases, eph and eck. Lambda-B1-Elk may be used in the production

XX of phosphorylated exogenous protein along with a further vector



CC encoding the desired exogenous protein. These plasmid may be used  
 CC to produce phosphorylated proteins in host cells which have no  
 CC intrinsic capacity for phosphorylation, eg. bacteria. The system  
 CC may be used for the expression of the phosphorylated kinase insert  
 CC domain of a growth factor receptor kinase eg. platelet-derived growth  
 CC factor receptor.

XX Sequence 1509 BP; 393 A; 374 C; 429 G; 313 T; 0 other;

#### Alignment Scores:

Pred. No.:	0.329	Length:	1509
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	14	Gaps:	0

US-09-397-967-16 (1-1099) x AA053470 (1-1509)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 409 GTGCACCGGACCTGCTGCTAGCAACATCTCTGCTG 444

#### RESULT 48

AA084888 standard; cDNA to mRNA; 1942 BP.

XX AA084888;

DT 26-OCT-1995 (first entry)

DE DNA encoding cytoplasmic tyrosine kinase.

KW cytoplasmic; tyrosine kinase; blood; cell differentiation;

KW screening; anticancer agent; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 208..1731

FT /tag= a

FT /product= cytoplasmic-tyrosine\_kinase

FT /note= "see AAR71133"

FT misc-feature 349..540

FT /tag= b

FT /note= "encodes SH3 domain (see AAR71129)"

FT misc-feature 571..795

FT /tag= c

FT /note= "encodes SH2 domain (see AAR71130)"

FT misc-feature 904..1641

FT /tag= d

FT /note= "encodes tyrosine kinase domain (see AAR71131)"

FT misc-feature 331..1728

FT /tag= e

FT /note= "encodes N-terminal truncated form of the enzyme (see AAR71132)"

XX W09506113-A.

XX 02-MAR-1995.

XX 25-AUG-1994; 94WO-JP01411.

XX 25-AUG-1993; 93JP-0210403.

XX 29-MAR-1994; 94JP-0038553.

XX (ASAH ) ASAMI KASEI KOGYO KK.

XX Sakano S;

XX WPI: 1995-106842/14.

XX P-PSDB: AAR71129-33.

DR

XX Cytoplasmic tyrosine kinase and antibody recognising it - for  
 PT screening chemical substances for tyrosine kinase inhibitory or  
 PT activating activity for use as cancer therapy

PS Claim 7; Page 49-50; 58pp; English.

CC This DNA encodes a cytoplasmic tyrosine kinase which has enhanced  
 CC expression in connection with blood cell differentiation. It was  
 CC isolated from the human HT-7 blood cell line. The DNA sequences and  
 CC antibodies raised against the enzyme, are useful for screening agents  
 CC for inhibiting or activating activity on the tyrosine kinase, for  
 CC use as anticancer agents.

XX Sequence 1942 BP; 365 A; 615 C; 651 G; 311 T; 0 other;

#### Alignment Scores:

Pred. No.:	0.416	Length:	1942
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AA084888 (1-1942)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

DB 1252 GTGCACCGGACCTGCTGCTAGCAACATCTCTGCTG 1287

#### RESULT 49

AAV44497 standard; cDNA; 1987 BP.

XX AAV44497;

DT 16-OCT-1998 (first entry)

DE Human matk cDNA.

KW Breast; cancer; matk; CSK homologous kinase; CHK; detection; diagnosis;

KW cytoplasmic protein; tyrosine kinase; EDB-2; negative regulator;

KW mitogenic signalling; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 263..1846

FT /tag= a

FT /product= CHK

FT /note= "CSK homologous kinase"

XX W09830704-A1.

XX 16-JUL-1998.

XX 07-JAN-1998; 98WO-US00420.

XX 16-JUN-1997; 97US-0876882.

XX 08-JAN-1997; 97US-0035228.

XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX Avraham H. Groopman JE;

XX WPI: 1998-399149/34.

XX P-PSDB: AAM64454.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PT Detecting breast cancer by detecting Csk homologous kinase  
 PT expression - especially in humans and use of Csk homologous kinase  
 PT in treatment or prophylaxis of breast cancer and for producing  
 PT medicaments

PS Disclosure; Fig 2; 54pp; English.

CC This sequence encodes a CSK homologous kinase (CHK) which is used in a  
 CC method of detecting cancer in breast tissue. The method allows diagnosis  
 CC of breast cancer in mammals, especially humans. It is based on the  
 CC discovery that a cytoplasmic protein tyrosine kinase, CHK, is expressed  
 CC in human breast tissue, but not in adjacent tissue. This protein can be  
 CC used to raise antibodies which can be included in compositions and  
 CC diagnostic kits for diagnosis of breast cancer. The presence of CHK in  
 CC breast tissue can also be determined using other standard methods (e.g.  
 CC Northern blotting) or by detecting nucleic acid sequences encoding all/a  
 CC portion of the protein (e.g. using hybridisation probes). Over-expression  
 CC of the receptor tyrosine kinase ErbB-2 has previously been associated  
 CC with the development of breast cancer, and CHK specifically interacts  
 CC with activated ErbB-2, and may function as a negative regulator of  
 CC ErbB-2 mediated mitogenic signalling. The compositions may also be used  
 CC to design drugs (e.g. which incorporate CHK analogues with greater  
 CC biological activity than CHK) and to identify CHK antagonists and  
 CC agonists for therapeutic use.

SQ Sequence 1987 BP; 369 A; 628 C; 672 G; 318 T; 0 other;

Alignment Scores:

Pred. No.:	0.425	Length:	1987
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	19	Gaps:	0

US-09-397-967-16 (1-1099) x AAV44497 (1-1987)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952  
 |||||||||||||||||||||||||||||||||||

DB 1307 GTGCACCGCGACTGGCGCGCACATCTGGTC 1342

RESULT 50  
 AAT00616  
 ID AAT00616 standard; cDNA; 2000 BP.

AC AAT00616;  
 XX  
 DT 26-MAR-1996 (first entry)  
 XX  
 DE Megakaryocyte kinase MKK1 cDNA.  
 XX  
 KW Megakaryocyte kinase-1; MKK1; cytoplasmic tyrosine kinase;  
 KW cellular signal transduction; leukaemia; myelosis; myelofibrosis;  
 KW gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 258..1781  
 FT /\*tag= a  
 XX  
 PN MO9529185-A1.  
 XX  
 PD 02-NOV-1995.  
 XX  
 PE 24-APR-1995; 95MO-US05008.  
 XX  
 PR 21-APR-1995; 95US-0426509.  
 PR 22-APR-1994; 94US-0232545.  
 XX  
 PA (ELAC) MAX PLANCK GES FÖRDERUNG WISSENSCHAFTEN.  
 PA (SUGEN-) SUGEN INC.  
 XX  
 PI Gishizky M, Sures I., Ullrich A;  
 XX  
 DR WPI: 1995-382959/49.  
 DR P-PSDB: AAR64181.  
 XX

PT New polynucleotide(s) encoding megakaryocyte tyrosine kinase(s) -  
 PT used to develop prods. for the treatment and diagnosis of kinase  
 PT related signal transduction abnormalities.

PS Claim 2; Fig 1A-C; 82pp; English.

CC Overlapping cDNA clones were combined to obtain a sequence (AAT00616)  
 CC coding for human megakaryocyte kinase MKK1 (AAR64181). The clones  
 CC were isolated from a fetal brain library using degenerate primers  
 CC (AAT00614-15) based on conserved regions within the kinase domain of  
 CC receptor tyrosine kinases. MKK polynucleotides can be used in the  
 CC prodn. in host cells of recombinant MKK, and in the gene therapy of  
 CC diseases such as acute megakaryocytic leukaemia, myelofibrosis and  
 CC acute megakaryocytic myelosis.

SQ Sequence 2000 BP; 375 A; 631 C; 674 G; 320 T; 0 other;

Alignment Scores:

Pred. No.:	0.428	Length:	2000
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAT00616 (1-2000)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952  
 |||||||||||||||||||||||||||||||||||

DB 1302 GTGCACCGCGACTGGCGCGCACATCTGGTC 1337

RESULT 51  
 AA241307  
 ID AA241307 standard; cDNA; 2170 BP.

AC AA241307;  
 XX  
 DT 18-JAN-2000 (first entry)  
 XX  
 DE Human normal ovarian tissue derived cDNA 86.  
 XX  
 KW Human; ovary; screening; ovarian cancer; treatment; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN DE19816395-A1.  
 XX  
 PD 07-OCT-1999.  
 XX  
 PE 03-APR-1998; 98DE-1016395.  
 XX  
 PR 03-APR-1998; 98DE-1016395.  
 XX  
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.  
 PA Rösenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
 PI WPI: 1999-552352/47.  
 DR  
 PT Nucleic acid sequences potentially useful in diagnosis or therapy of  
 PT ovarian cancer  
 XX  
 PS Claim 3; Page 196; 274pp; German.

CC This invention describes novel nucleic acid sequences that are highly  
 CC expressed in normal ovary tissue. Artificial chromosomes and cosmid  
 CC clones containing the sequences can be used as gene transfer vehicles.  
 CC The sequences can be used to produce DNA fragments containing  
 CC full-length genes. Host cells transformed with the sequences can be used  
 CC to produce polypeptides or polypeptide fragments, which can be used to  
 CC screen phage displays for polypeptides that bind to them, or as tools for  
 CC identifying agents active against ovarian cancer, or to prepare  
 CC medicaments for treating ovarian cancer. The cDNA sequences can be used



CC stability and thus intracellular concentration.  
 XX Sequence 2784 BP; 770 A; 650 C; 731 G; 633 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 0.584 Length: 2784  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x ABL91661 (1-2784)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952  
 |||||||||||||||||||||||||||||||||||  
 Db 2050 GTGCATCGTATCTGCGCCGACGAGACATCTGTG 2085

RESULT 54  
 AAQ90654  
 ID AAQ90654 standard; cDNA; 2820 BP.  
 AC  
 XX AAQ90654;  
 XX  
 DT 11-NOV-1995 (first entry)  
 XX  
 DE Eph-related PTK Cdk8 CDNA.  
 XX  
 KW Cdk8; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;  
 KM prognosis; ss.  
 XX  
 OS Gallus sp.  
 XX  
 FH Location/Qualifiers  
 FT 2..2551  
 FT CDS /\*tag- a

XX  
 PN MO9515375-A.  
 XX  
 PD 08-JUN-1995.  
 XX  
 PF 07-SEP-1994; 94MO-US10140.  
 XX  
 PR 03-DEC-1993; 93US-0162809.  
 XX  
 PA (LJOL-) LA JOLLA CANCER RES FOUND.  
 XX  
 PI Pasquale EB, Sajjadi FG;  
 XX  
 DR WPI: 1995-215256/28.  
 DR P-PSDB: AAR75706.  
 XX  
 PT Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing  
 PT cancer.  
 XX  
 PS Disclosure: Page 50-53; 129pp; English.  
 XX  
 CC The chick Eph-related PTK cDNA clone given in AAQ90653 encoded Cdk8  
 CC (AAR75706) that showed 74% amino acid identity in the catalytic  
 CC and C-terminal regions to Cdk5 (AAR75712) and novel Cdk6 (AAR75704),  
 CC Cdk7 (AAR75705) and Cdk10 (AAR75708). Cdk8 was expressed in adult  
 CC brain and retina and at lower levels in kidney, lung, skeletal  
 CC muscle and thymus.  
 XX  
 SO Sequence 2820 BP; 753 A; 693 C; 731 G; 643 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.591 Length: 2820  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 16 Gaps: 0

US-09-397-967-16 (1-1099) x AAQ90654 (1-2820)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952  
 |||||||||||||||||||||||||||||||||||  
 Db 1817 GTGCATCGGATCTGCTGCTCGAACAACATCTGTG 1852

RESULT 55  
 ABL91658  
 ID ABL91658 standard; DNA; 2955 BP.  
 AC  
 XX ABL91658;  
 XX  
 DT 28-MAY-2002 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 1.  
 XX  
 KW Human; HIV; HCV; gene expression; oligonucleotide; tumour; pathogen;  
 KM Plasmidium; virus; viroid; cytokine; prion; antisense oligonucleotide;  
 KW cytosolic; virucide; protozoacide; antibacterial; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN DE10100586-C1.  
 XX  
 PD 11-APR-2002.  
 XX  
 PE 09-JAN-2001; 2001DE-1000586.  
 XX  
 PR 09-JAN-2001; 2001DE-1000586.  
 XX  
 PA (RIBO-) RIBOPHARMA AG.  
 XX  
 PI Kreutzer R, Lämmer S, Rost S, Hadwiger P;  
 PI WPI: 2002-270454/32.  
 DR  
 XX  
 PT Inhibiting gene expression in cells, useful for e.g. treating tumors,  
 PT by introducing double-stranded complementary oligoRNA having unpaired  
 PT terminal bases  
 XX  
 PS Claim 13; Page 5-6; 104pp; German.  
 XX  
 CC The invention relates to a method for inhibiting expression of a target  
 CC gene (ABL91658-ABL91797) in a cell by introducing at least one  
 CC oligonucleotide that has a double-stranded structure consisting of at  
 CC most 49 sequential nucleotide pairs, with at least part of one strand  
 CC complementary with the target gene and has at least one end a  
 CC single-stranded segment of 1-4 nt. The method provides  
 CC oligonucleotides for antisense inhibition of gene expression useful  
 CC e.g. for treating tumours but the oligonucleotides may also be  
 CC directed against genes present in pathogens (e.g. plasmidium or  
 CC viruses/viroids), pathogenic on humans, animals or plants) or against  
 CC cytokine, id, developmental or prion genes. The method provides more  
 CC effective inhibition of gene expression than use of known  
 CC oligonucleotides, probably because the unpaired overhang increases  
 CC stability and thus intracellular concentration.  
 XX  
 SO Sequence 2955 BP; 597 A; 868 C; 904 G; 586 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.617 Length: 2955  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x ABL91658 (1-2955)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952  
 |||||||||||||||||||||||||||||||||||  
 Db 2260 GTCCACCGGACCTGCTGCTCCAGAAACATCTTGTG 2295



KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;  
 KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;  
 KW Cystostatic; virucide; protozoacide; antibacterial; ds.  
 OS Homo sapiens.  
 XX DE10100586-CL.  
 PN 11-APR-2002.  
 PD 09-JAN-2001; 2001DE-1000586.  
 PF 09-JAN-2001; 2001DE-1000586.  
 PR 09-JAN-2001; 2001DE-1000586.  
 PA (RIBO-) RIBOPHARMA AG.  
 PI Kreutzer R, Limmer S, Rost S, Hadwiger P;  
 DR WPI: 2002-270454/32.  
 XX Inhibiting gene expression in cells, useful for e.g. treating tumors,  
 PT by introducing double-stranded complementary oligoRNA having unpaired  
 PT terminal bases  
 XX Claim 13; Page 24-25; 104pp; German.  
 PS  
 XX The invention relates to a method for inhibiting expression of a target  
 CC gene (ABL91658-ABL91797) in a cell by introducing at least one  
 CC oligoribonucleotide that has a double-stranded structure consisting of at  
 CC most 49 sequential nucleotide pairs, with at least part of one strand  
 CC complementary with the target gene and has at least one end a  
 CC single-stranded segment of 1-4 nt. The method provides  
 CC oligoribonucleotides for antisense inhibition of gene expression useful  
 CC e.g. for treating tumours but the oligoribonucleotides may also be  
 CC directed against genes present in pathogens (e.g. Plasmodium or  
 CC viruses/viroids, pathogenic on humans, animals or plants) or against  
 CC cytokine, id. developmental or prion genes. The method provides more  
 CC effective inhibition of gene expression than use of known  
 CC oligonucleotides, probably because the unpaired overhang increases  
 CC stability and thus intracellular concentration.  
 CC  
 SQ Sequence 2964 BP; 563 A; 917 C; 940 G; 544 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.619 Length: 2964  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-397-967-16 (1-1099) x ABL91681 (1-2964)  
 QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 ||||||||||||||||||||||||||||||||  
 DB 2209 GTCCACCGAGACTGGCTGCTCGCAACATCTTAGTC 2244  
 RESULT 59  
 ID AAG090971 standard; cDNA to mRNA; 2982 BP.  
 XX  
 AC AAG090971;  
 XX  
 DT 24-NOV-1995 (first entry)  
 XX  
 DE Protein p140 cDNA from rat skeletal muscle myoblast cell line L6.  
 XX  
 KW Protein p140; insulin; tyrosine phosphorylation; ss.  
 OS Rattus rattus.  
 XX  
 PN Fp659883-A.  
 XX

PD 28-JUN-1995.  
 XX  
 XX 24-NOV-1994; 94EP-0118524.  
 PF  
 XX 24-NOV-1993; 93JP-0315806.  
 PR  
 XX (ONCV ) ONO PHARM CO LTD.  
 PA  
 XX Kitagawa K, Ohno H, Tajima H;  
 PI  
 XX WPI: 1995-226291/30.  
 DR  
 XX Isolated protein p140 polypeptide - and treatment of diabetes based  
 PT on tyrosine phosphorylation of protein p140.  
 PT  
 XX Claim 4; Page 23-25; 42pp; English.  
 PS  
 CC p140 is used for the prevention and treatment of diabetes. Dosage  
 CC is 10 microg-1000 mg (p.o.) or 10 microg-100mg (i.v.). A  
 CC polypeptide of protein p140 having the sequence in AAR75843 is  
 CC claimed. also claimed is a DNA encoding it, specifically,  
 CC AAG090971 and AAG090972.  
 CC  
 SQ Sequence 2982 BP; 639 A; 889 C; 853 G; 601 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.623 Length: 2982  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 16 Gaps: 0  
 US-09-397-967-16 (1-1099) x AAG090971 (1-2982)  
 QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 ||||||||||||||||||||||||||||||||  
 DB 2248 GTGCACCGTAGCTGCTGCGCCGCAACATCTTAGTC 2283  
 RESULT 60  
 ID ABL91662 standard; DNA; 2997 BP.  
 XX  
 AC ABL91662;  
 XX  
 DT 28-MAY-2002 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 5.  
 XX  
 KW Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;  
 KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;  
 KW Cystostatic; virucide; protozoacide; antibacterial; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN DE10100586-CL.  
 PD 11-APR-2002.  
 PD  
 PF 09-JAN-2001; 2001DE-1000586.  
 PF  
 PR 09-JAN-2001; 2001DE-1000586.  
 PR  
 PA (RIBO-) RIBOPHARMA AG.  
 PA  
 PI Kreutzer R, Limmer S, Rost S, Hadwiger P;  
 PI  
 DR WPI: 2002-270454/32.  
 DR  
 XX Inhibiting gene expression in cells, useful for e.g. treating tumors,  
 PT by introducing double-stranded complementary oligoRNA having unpaired  
 PT terminal bases  
 XX

PS Claim 13: Page 9-10; 104pp; German.  
XX  
CC The invention relates to a method for inhibiting expression of a target  
CC gene (ABL91658-ABL91797) in a cell by introducing at least one  
CC oligoribonucleotide that has a double-stranded structure consisting of at  
CC most 49 sequential nucleotide pairs, with at least part of one strand  
CC complementary with the target gene and has at least one end a  
CC single-stranded segment of 1-4 nt. The method provides  
CC oligoribonucleotides for antisense inhibition of gene expression useful  
CC e.g. for treating tumours but the oligoribonucleotides may also be  
CC directed against genes present in pathogens (e.g. Plasmodium or  
CC viruses/viroids, pathogenic on humans, animals or plants) or against  
CC cytokine, id, developmental or prion genes. The method provides more  
CC effective inhibition of gene expression than use of known  
CC oligonucleotides, probably because the unpaired overhang increases  
CC stability and thus intracellular concentration.  
SQ Sequence 2997 BP; 901 A; 593 C; 732 G; 771 T; 0 other:  
Alignment Scores:  
Pred. No.: 0.625 Length: 2997  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 24 Gaps: 0  
US-09-397-967-16 (1-1099) x ABL91662 (1-2997)  
OY 941 VALHISARGSPLEUALAALARGASNIIELEUVAL 952  
ID ABL91680 standard; DNA: 2997 BP.  
XX ABL91680;  
AC ABL91680;  
XX 28-MAY-2002 (first entry)  
DT  
DT  
DT  
DE Human polynucleotide SEQ ID NO 23.  
XX  
XX  
XX Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;  
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;  
KW cytostatic; virucide; protozoicide; antibacterial; ds.  
XX  
OS Homo sapiens.  
OS  
XX DE10100586-C1.  
PN  
XX 11-APR-2002.  
PD  
XX 09-JAN-2001; 2001DE-1000586.  
PF  
XX 09-JAN-2001; 2001DE-1000586.  
PR  
XX 09-JAN-2001; 2001DE-1000586.  
XX  
XX (RIBO-) RIBOPHARMA AG.  
PA  
XX Kreutzer R, Limmer S, Rost S, Hadwiger P;  
PI  
XX WPI: 2002-270454/32.  
DR  
XX  
XX Inhibiting gene expression in cells, useful for e.g. treating tumours,  
PT by introducing double-stranded complementary oligoRNA having unpaired  
PT terminal bases -  
XX  
XX  
XX Claim 13: Page 23-24; 104pp; German.  
PS  
XX The invention relates to a method for inhibiting expression of a target  
CC gene (ABL91658-ABL91797) in a cell by introducing at least one  
CC oligoribonucleotide that has a double-stranded structure consisting of at  
CC most 49 sequential nucleotide pairs, with at least part of one strand

CC complementary with the target gene and has at least one end a  
CC single-stranded segment of 1-4 nt. The method provides  
CC oligoribonucleotides for antisense inhibition of gene expression useful  
CC e.g. for treating tumours but the oligoribonucleotides may also be  
CC directed against genes present in pathogens (e.g. Plasmodium or  
CC viruses/viroids, pathogenic on humans, animals or plants) or against  
CC cytokine, id, developmental or prion genes. The method provides more  
CC effective inhibition of gene expression than use of known  
CC oligonucleotides, probably because the unpaired overhang increases  
CC stability and thus intracellular concentration.  
SQ Sequence 2997 BP; 590 A; 940 C; 875 G; 592 T; 0 other:  
Alignment Scores:  
Pred. No.: 0.625 Length: 2997  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 24 Gaps: 0  
US-09-397-967-16 (1-1099) x ABL91680 (1-2997)  
OY 941 VALHISARGSPLEUALAALARGASNIIELEUVAL 952  
ID ABL91659 standard; DNA: 3042 BP.  
XX ABL91659;  
AC ABL91659;  
XX 28-MAY-2002 (first entry)  
DT  
DT  
DT  
DE Human polynucleotide SEQ ID NO 2.  
XX  
XX  
XX Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;  
KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;  
KW cytostatic; virucide; protozoicide; antibacterial; ds.  
XX  
OS Homo sapiens.  
OS  
XX DE10100586-C1.  
PN  
XX 11-APR-2002.  
PD  
XX 09-JAN-2001; 2001DE-1000586.  
PF  
XX 09-JAN-2001; 2001DE-1000586.  
PR  
XX 09-JAN-2001; 2001DE-1000586.  
XX  
XX (RIBO-) RIBOPHARMA AG.  
PA  
XX Kreutzer R, Limmer S, Rost S, Hadwiger P;  
PI  
XX WPI: 2002-270454/32.  
DR  
XX  
XX Inhibiting gene expression in cells, useful for e.g. treating tumours,  
PT by introducing double-stranded complementary oligoRNA having unpaired  
PT terminal bases -  
XX  
XX  
XX Claim 13: Page 6-7; 104pp; German.  
PS  
XX The invention relates to a method for inhibiting expression of a target  
CC gene (ABL91658-ABL91797) in a cell by introducing at least one  
CC oligoribonucleotide that has a double-stranded structure consisting of at  
CC most 49 sequential nucleotide pairs, with at least part of one strand  
CC complementary with the target gene and has at least one end a  
CC single-stranded segment of 1-4 nt. The method provides  
CC oligoribonucleotides for antisense inhibition of gene expression useful  
CC e.g. for treating tumours but the oligoribonucleotides may also be  
CC directed against genes present in pathogens (e.g. Plasmodium or  
CC viruses/viroids, pathogenic on humans, animals or plants) or against

CC cytokine, Id, developmental or prion genes. The method provides more  
 CC effective inhibition of gene expression than use of known  
 CC oligonucleotides, probably because the unpaired overhang increases  
 CC stability and thus intracellular concentration.

XX Sequence 3042 BP; 600 A; 941 C; 960 G; 541 T; 0 other;

#### Alignment Scores:

Pred. No.:	0.634	Length:	3042
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	24	Gaps:	0

US-09-397-967-16 (1-1099) x ABL91659 (1-3042)

Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 2317 GTGCACCGTGCCTGCTGCGCAACATCCTCGTC 2352

#### RESULT 63

ID AAT07308 standard; CDNA: 3105 BP.

XX AAT07308;

DT 19-MAR-1996 (first entry)

XX Receptor tyrosine kinase (neural kinase) cDNA.

XX Receptor tyrosine kinase; neural kinase; Nuk gene; axon;

XX axonogenesis; nerve disorder; gene therapy; transgenic animal; ss.

XX Mus musculus.

XX Key Location/Qualifiers

XX CDS 1..2985

XX sig\_peptide /tag= a

XX mat\_peptide 79..2982

XX /tag= c

XX WO9530326-A1.

XX 09-NOV-1995.

XX 28-APR-1995; 95WO-CA00254.

XX 29-APR-1994; 94US-0235407.

XX (MOUN ) MOUNT SINAI HOSPITAL CORP.

XX Henkemeyer M, Letwin K, Pawson A;

XX WPI; 1995-393299/50.

XX P-PSDB; AAR87018.

XX DNA encoding neural receptor tyrosine kinase - useful in gene

XX therapy of nerve disorders, and for diagnosis and identification of

XX therapeutic agents

XX Claim 2; Page 68-70; 103pp; English.

XX CDNA clones pNukRACE A2 and K2 were combined to obtain a sequence

XX (AAT07308) coding for a novel receptor tyrosine kinase, designated

XX neural kinase (Nuk) (AAR87018). The clones were obtd. from a cDNA

XX library of mouse embryo cDNA in lambda-gt10 probed with a partial Nuk

XX cDNA insert. The gene was mapped to the distal end of chromosome 4

XX near the and-1 mutation. The cDNA is used to produce recombinant

XX Nuk or transgenic animal models for studies of Nuk function, or as

XX probes to detect genes or diagnose conditions associated with Nuk

CC gene mutation, or in gene (antisense) therapy of disorders of the  
 CC nervous system.

XX Sequence 3105 BP; 711 A; 937 C; 847 G; 610 T; 0 other;

#### Alignment Scores:

Pred. No.:	0.647	Length:	3105
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAT07308 (1-3105)

Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 2251 GTGCACCGTGCCTGCTGCGAACAATCCTCGTC 2286

#### RESULT 64

ID AAT84528 standard; CDNA: 3105 BP.

XX AAT84528;

DT 02-DEC-1997 (first entry)

XX Mouse Nuk tyrosine kinase cDNA.

XX Nuk tyrosine kinase; Eph receptor tyrosine kinase;

XX signal transduction; axonogenesis; neurodegenerative disease;

XX Alzheimer's disease; Parkinson's disease; Huntington's disease;

XX multiple sclerosis; amyotrophic lateral sclerosis;

XX Wernicke's disease; nerve damage; trauma; ischemia; stroke; ss.

XX Mus musculus.

XX Key Location/Qualifiers

XX CDS 1..2985

XX sig\_peptide /tag= a

XX mat\_peptide 79..2982

XX /tag= c

XX WO9714966-A1.

XX 24-APR-1997.

XX 10-OCT-1996; 96WO-CA00679.

XX 13-OCT-1995; 95US-0005518.

XX (MOUN ) MOUNT SINAI HOSPITAL CORP.

XX Henkemeyer M, Pawson A;

XX WPI; 1997-245245/22.

XX P-PSDB; AAM26366.

XX Activation of ligand regulatory pathways by Eph subfamily receptor

XX tyrosine kinases - for stimulating or inhibiting axonogenesis,

XX useful for treatment of e.g. neurodegenerative diseases such as

XX Alzheimer's or Parkinson's diseases

XX Disclosure; Page 24-25; 55pp; English.

XX This cDNA sequence codes for murine Nuk tyrosine kinase (AAM26366),

XX an Eph subfamily receptor tyrosine kinase essential for formation

XX of the medial tract of the anterior commissure of the brain. Its

XX sequence was deduced from clones isolated from an embryo cDNA

XX library in lambda gt10. The Nuk gene maps to the distal end of

XX chromosome 4 near the and-1 mutation. Eph subfamily receptor



CC tyrosine kinases (e.g. Nuk extracellular domain polypeptides)  
 CC can be used in claimed methods to: activate a ligand regulatory  
 CC pathway in a cell; identify a substance able to bind a ligand for  
 CC an Eph subfamily receptor tyrosine kinase; and to affect neuronal  
 CC development or regeneration, especially the stimulation or  
 CC inhibition of axonogenesis, in a mammal. Activation of the ligand  
 CC regulatory pathways results in downstream activation of a series of  
 CC division, cytoskeletal architecture, cell metabolism, cell  
 CC migration and cell-cell interactions. Substances which activate  
 CC the ligand regulatory pathway may be used for stimulating or  
 CC inhibiting neuronal development, regeneration and axonal migration  
 CC associated with neurodegenerative disease e.g. Alzheimer's,  
 CC Parkinson's or Huntington's diseases, multiple sclerosis,  
 CC amyotrophic lateral sclerosis, deficiency diseases such as  
 CC Wernicke's disease, peripheral nerve damage, trauma and ischemia  
 CC resulting from stroke.

SO Sequence 3105 BP; 710 A; 937 C; 848 G; 610 T; 0 other;

# Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0.647	12.00	100.00%	100.00%	1.09%	3105	12	0	0	0	0

US-09-397-967-16 (1-1099) x AAT84538 (1-3105)

OY 941 VALHISARGASPLEUALAIALARGASNIJLEUVAL 952

DB 2251 GTGCACCGTGACCTGCTGCTGGAACATCTCTGTC 2286

RESULT 65

AAS03812

ID AAS03812 standard; cDNA: 3105 BP.

XX AAS03812;

XX 29-AUG-2001 (first entry)

DE Murine neural kinase (Nuk) cDNA.

KW Neural kinase; Nuk; receptor tyrosine kinase; axonal migration; stroke;  
 KW nerve fibre; cell-cell interaction; axonogenesis; neuronal development;  
 KW regeneration; neurodegenerative disorder; Alzheimer's disease; ischemia;  
 KW Parkinson's disease; Huntington's disease; demyelinating disease; ss;  
 KW multiple sclerosis; amyotrophic lateral sclerosis; deficiency disease;  
 KW Wernicke's disease; nutritional polyneuropathy; multistem degeneration;  
 KW progressive supranuclear palsy; Shy Drager's syndrome; mouse;  
 KW olivoponto cerebellar atrophy; peripheral nerve damage.

XX Mus musculus.

XX Key Location/Qualifiers

XX CDS

XX 1..2985

XX /tag- a

XX /product- "Murine neural kinase"

XX 1..78

XX /tag- b

XX 79..2982

XX /tag- c

XX /product- "Murine neural kinase"

XX US6218356-B1.

XX 17-APR-2001.

XX 13-OCT-1995; 95US-0542635.

XX 29-APR-1994; 94US-0235407.

XX 28-APR-1995; 95WO-CA00254.

XX (MOUN ) MOUNT SINAI HOSPITAL CORP.

XX Pawsen A, Henkemeyer M, Letwin K;

XX WPI: 2001-289845/30.

XX P-PSDB; AAU01907.

XX New composition comprising neural receptor tyrosine kinase protein  
 XX useful for inhibiting or stimulating axonogenesis, neuronal  
 XX development, or regeneration and axonal migration

XX Disclosure: Fig 1; 86pp; English.

XX The sequence represents a cDNA molecule which encodes a mouse neural  
 XX kinase (Nuk) polypeptide. The murine Nuk locus has been mapped to the  
 XX distal end of chromosome four. The polypeptide is a novel receptor  
 XX tyrosine kinase protein, and is found to be expressed at high levels  
 XX within migrating axons and is associated with nerve fibres. It functions  
 XX to regulate specific cell-cell interactions during early development of  
 XX the nervous system and in axonogenesis. Substances which bind to the Nuk  
 XX neuronal development, regeneration and axonal migration associated with  
 XX neurodegenerative disorders and conditions involving trauma and injury to  
 XX the nervous system. These disorders include Alzheimer's disease,  
 XX Parkinson's disease, Huntington's disease, demyelinating diseases such as  
 XX multiple sclerosis, amyotrophic lateral sclerosis, deficiency diseases  
 XX such as Wernicke's disease and nutritional polyneuropathy, progressive  
 XX supranuclear palsy, Shy Drager's syndrome, multistem degeneration, olivop  
 XX ponto cerebellar atrophy, peripheral nerve damage, and ischemia  
 XX resulting from stroke. The proteins may be used to prepare antibodies  
 XX having specificity for Nuk proteins, which can be used to diagnose or  
 XX treat disorders of the nervous system. These proteins are also used for  
 XX screening agonists or antagonists of the interactions of the Nuk proteins  
 XX with binding molecules.

XX Sequence 3105 BP; 710 A; 937 C; 848 G; 610 T; 0 other;

# Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0.647	12.00	100.00%	100.00%	1.09%	3105	12	0	0	0	0

US-09-397-967-16 (1-1099) x AAS03812 (1-3105)

OY 941 VALHISARGASPLEUALAIALARGASNIJLEUVAL 952

DB 2251 GTGCACCGTGACCTGCTGCTGGAACATCTCTGTC 2286

RESULT 66

AAS69876

ID AAS69876 standard; cDNA: 3107 BP.

XX AAS69876;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #5680.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 PA (HXSE-) HXSEQ INC.  
 PI Drmanac RT, Liu C, Tang YT;  
 XX WPI: 2001-639362/73.  
 DR P-PSDB; ABG05689.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 PS  
 PS Claim 1; SEQ ID NO 5680; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AA564197-AA594564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 XX Sequence 3107 BP; 850 A; 720 C; 812 G; 725 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 0.647 Length: 3107  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 23 Gaps: 0  
 US-09-397-967-16 (1-1099) x AA569876 (1-3107)  
 OY 941 VALHISARGASPLEUAAIAAARGASNIILEUVAL 952  
 DB 2260 GTGCATCGATCTGCGCCGACGGAACATCTGTGTG 2295  
 RESULT 67  
 ABL65818  
 ID ABL65818 standard; DNA; 3107 BP.  
 XX  
 AC ABL65818;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Lung cancer related gene sequence SEQ ID NO:4155.  
 XX  
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
 KW gene; ds.  
 XX  
 XX Homo sapiens.  
 OS  
 PM W0200194629-A2.

XX 13-DEC-2001.  
 PD 30-MAY-2001; 2001WO-US10838.  
 XX  
 PF 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-233133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 22-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 26-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237112P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 02-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237596P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 PA (AVAL-) AVALON PHARM.  
 XX  
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX WPI: 2002-188264/24.  
 DR  
 XX  
 XX Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set  
 XX  
 PS Claim 1; SEQ ID 4155; 44pp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC result of M1, and the data is sufficient to convey the chemical  
 CC structure and/or properties of the agent. M1 can be used in the

CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilms' tumour.  
 XX

SO Sequence 3107 BP; 850 A; 720 C; 812 G; 725 T; 0 other;

#### Alignment Scores:

Pred. No.:	0.647	Length:	3107
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	24	Gaps:	0

US-09-397-967-16 (1-1099) x ABL5818 (1-3107)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952

DB 2260 GTGCATCGTGCATCTGCCGACGAGACATCTCTGTG 2295

#### RESULT 68

ID AAT02948 standard; cDNA; 3116 BP.

XX AAT02948;

DT 16-APR-1996 (first entry)

DE Eph-like receptor protein tyrosine kinase HEK8 CDNA.

KW Eph-like receptor protein tyrosine kinase; PTK; HEK8;

KM human eph-like kinase; therapy; diagnosis; antibody; vector; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 34..2994

XX /\*tag= a

PN MO528484-A1.

XX 26-OCT-1995.

PF 14-APR-1995; 95WO-US04681.

PR 15-APR-1994; 94US-0229509.

XX (AMGE-) AMGEN INC.

PI Fox GM, Jing S, Welcher AA;

DR WPI: 1995-373799/48.

DR P-PSDB: AAR85091.

PT New nucleic acid encoding Eph-like receptor tyrosine kinase(s)  
 PT and related vectors, host cells, proteins, antibodies etc., used  
 PT diagnostically and therapeutically to modulate receptor activation  
 or prodn.

XX Claim 1; Page 57-62; 133pp; English.

CC CDNAs (AAT02946-49) coding for 4 novel human Eph-like receptor protein  
 CC tyrosine kinases, HEK5, HEK7, HEK8 and HEK11 (AAR85089-92),  
 CC respectively, were isolated from a human foetal brain cDNA library using  
 CC a directed PCR approach with primers (see AAT02960-61) based on conserved  
 CC regions of receptor PTKs and Eph-like receptor PTKs. HEK5, HEK7 and HEK8  
 CC show extensive homology to the catalytic domain of chicken Eph-like  
 CC receptors Cek5, Cek7 and Cek8. HEK11 shows no homology to any known  
 CC Eph-like receptor. The isolated cDNAs are used for prodn. of  
 CC recombinant HEKs and chimeric receptors, in hybridisation assays, and  
 CC to detect abnormalities in HEK receptor genes.

XX  
 SO Sequence 3116 BP; 859 A; 720 C; 812 G; 725 T; 0 other;

#### Alignment Scores:

Pred. No.:	0.649	Length:	3116
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAT02948 (1-3116)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952

DB 2260 GTGCATCGTGCATCTGCCGACGAGACATCTCTGTG 2295

#### RESULT 69

ID AAG90652 standard; cDNA; 3133 BP.

XX AAG90652;

DT 11-NOV-1995 (first entry)

DE Eph-related tyrosine kinase CEK6 CDNA.

KW Cek6; Eph: protein tyrosine-kinase; PTK; cancer; diagnosis;

KM prognosis; ss.

OS Gallus sp.

XX Key

FT CDS 3..419

FT CDS /\*tag= a

XX /\*tag= b

PN MO9515375-A.

XX 08-JUN-1995.

PF 07-SEP-1994; 94WO-US10140.

PR 03-DEC-1993; 93US-0162809.

XX (LJOL-) LA JOLLA CANCER RES FOUND.

PI Pasquale EB, Sajjadi FG;

DR WPI: 1995-215256/28.

DR P-PSDB: AAR75704.

PT Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing  
 PT cancer.  
 PT Disclosure: Page 37-41; 129pp; English.

CC Novel Eph-related PTK Cek6 cDNA clones (AAG90652) were isolated from  
 CC chick embryo and embryonic brain cDNA libraries in phage lambda gt11.  
 CC The encoded Cek6 protein (AAR75704) is closely related to rat Elk,  
 CC Ceks (AAR75712) and Cek10 (AAR75708). Cek6 transcripts were found in  
 CC 10-day embryos and in adult brain, lung, heart and skeletal muscle.

SO Sequence 3133 BP; 718 A; 918 C; 922 G; 575 T; 0 other;

#### Alignment Scores:

Pred. No.:	0.652	Length:	3133
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AA090652 (1-3133)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 |||||||  
 Db 2125 GTCACAGGATCTGCGCGCCAGAACATCTCTGTC 2160

RESULT 70

AAA09322 standard; DNA: 3151 BP.

Id AAA09322:

AC AAA09322:

DT 10-AUG-2000 (first entry)

DE Human cancer associated antigen precursor DNA, clone NY-REN-47.

XX renal cancer; cancer associated antigen precursor; diagnosis;

KW cytosolic; ERK tyrosine kinase; ss.

XX Homo sapiens.

OS WO200020587-A2.

PN 13-APR-2000.

PD 04-OCT-1999; 99WO-US22873.

PF 05-OCT-1998; 98US-0166300.

PR 05-OCT-1998; 98US-0166350.

XX (LUDWIG) LUDWIG INST CANCER RES.

XX Obata Y, Gout I, Tureci O, Sahin U, Pfeunedschuh M, Scanlan MJ;

PI Stockert E, Chen Y, Old LJ, Jager E, Knuth A;

PT abnormal expression of human cancer associated antigens

PS Claim 57, Page 93-94; 121pp; English.

XX AAA09321.45 were isolated by SEREX screening from a renal cancer

CC cell line 1973/10.4. Homology searching revealed that these clones

CC correspond to known genes. The present sequence has identity with the

CC ERK tyrosine kinase gene. The genes encode cancer associated antigen

CC precursors. These gene products are useful in methods for preventing,

CC diagnosing and/or treating disorders, especially cancer, associated with

CC abnormal expression of a human cancer associated antigen. The method

CC specifically binds to the nucleic acid molecule or expression product

CC (or fragment) complexed with a human leukocyte antigen (HLA) molecule

CC and determining the interaction between the agent and the nucleic acid

CC molecule or the expression product as a determination of the disorder.

XX Sequence 3151 BP; 707 A; 974 C; 874 G; 596 T; 0 other;

US-09-397-967-16 (1-1099) x AAA09322 (1-3151)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 2236 GTTCACCGTACCTGCTGCCGCAACATCTCTGTC 2271

RESULT 71

ABL91679 standard; DNA: 3168 BP.

XX ABL91679:

DT 28-MAY-2002 (first entry)

DE Human polynucleotide SEQ ID NO 22.

XX Human; HIV; HCV; gene expression; oligonucleotide; tumour; pathogen;

KW Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;

KW cytosolic; virucide; protozoacide; antibacterial; ds.

XX Homo sapiens.

OS DE10100586-C1.

PN 11-APR-2002.

PD 09-JAN-2001; 2001DE-1000586.

PF 09-JAN-2001; 2001DE-1000586.

PR (RIBO-) RIBOPHARMA AG.

XX Kreutzler R, Limmer S, Rost S, Hadwiger P;

PI WPI; 2002-270454/32.

XX Inhibiting gene expression in cells, useful for e.g. treating tumors,

PT by introducing double-stranded complementary oligoRNA having unpaired

PT terminal bases

PS Claim 13, Page 22-23; 104pp; German.

XX The invention relates to a method for inhibiting expression of a target

CC gene (ABL91658-ABL91797) in a cell by introducing at least one

CC oligonucleotide that has a double-stranded structure consisting of at

CC most 49 sequential nucleotide pairs, with at least one end a

CC complementary with the target gene and has at least one end a

CC single-stranded segment of 1-4 nt. The method provides

CC oligonucleotides for antisense inhibition of gene expression useful

CC e.g. for treating tumors but the oligonucleotides may also be

CC directed against genes present in pathogens (e.g. Plasmodium or

CC viruses/viroids, pathogenic on humans, animals or plants) or against

CC cytokine. Id. development of gene expression than use of known

CC oligonucleotides, probably because the unpaired overhang increases

CC stability and thus intracellular concentration.

XX Sequence 3168 BP; 732 A; 947 C; 902 G; 587 T; 0 other;

US-09-397-967-16 (1-1099) x ABL91679 (1-3168)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 2227 GTTCACCGTACCTGCTGCCGCAACATCTCTGTC 2262

RESULT 72

AA87118 standard; CDNA: 3342 BP.

DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #22922.  
 DE Human: chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 KW Homo sapiens.  
 XX MO200175067-A2.  
 XX 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US08631.  
 XX 31-MAR-2000; 2000US-0540217.  
 XX 23-AUG-2000; 2000US-0649167.  
 XX (HWE-) HWEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 PI WPI: 2001-639362/73.  
 DR P-PSDB: ABG22931.  
 XX New isolated polynucleotide and encoded polypeptides; useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 1: SEQ ID No 22922: 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 3342 BP; 898 A; 796 C; 867 G; 781 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0.693 Length: 3342  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 23 Gaps: 0  
 US-09-397-967-16 (1-1099) x AAS87118 (1-3342)  
 QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 DB 2494 GTGCATCGTATCTGGCCGACGAGACATCTGTG 2529  
 RESULT 73  
 AAT03100  
 ID AAT03100 standard; DNA: 3348 BP.  
 XX

AC AAT03100;  
 XX 14-FEB-1996 (first entry).  
 DE Protein tyrosine-kinase bptk7 gene.  
 XX Protein tyrosine-kinase bptk7; agonist; cell growth;  
 KW differentialiation; ss.  
 XX Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..2961  
 FT sig\_peptide /\*tag= a  
 FT mat\_peptide /\*tag= b  
 FT /\*tag= c  
 XX  
 PN WO9527061-A1.  
 XX 12-OCT-1995.  
 XX 04-APR-1995; 95WO-US04228.  
 XX 04-APR-1994; 94US-0222616.  
 XX (GETH ) GENENTECH INC.  
 XX  
 PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;  
 PI Wood WI;  
 DR WPI: 1995-366160/47.  
 DR P-PSDB: AAR85936.  
 XX  
 PT Agonist antibodies which activate specific protein tyrosine  
 PT kinase(s) - also activate chimeric proteins of kinase extracellular  
 PT domain and Ig constant domain, useful for studying, and therapeutic  
 PT modulation of, cell growth and differentiation  
 XX  
 PS Disclosure: Page 88-92; 125pp; English.  
 XX  
 CC DNA probes based on protein tyrosine-kinase (PTK) sequences were used  
 CC to screen cDNA libraries to identify novel PTK genes. The bptks,  
 CC bptk1, bptk2, bptk3, bptk4, bptk5 and bptk7 (AAR85924-28 and AAR85935,  
 CC respectively) are expressed in human brain tissue and show homology  
 CC to known PTKs. A full-length sequence for the bptk7 gene (AAT03100)  
 CC was obt'd. This gene may be used to design new drugs, peptides and  
 CC antisense constructs that modulate PTK activity.  
 XX  
 SQ Sequence 3348 BP; 924 A; 769 C; 855 G; 800 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0.694 Length: 3348  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 16 Gaps: 0  
 US-09-397-967-16 (1-1099) x AAT03100 (1-3348)  
 QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 DB 2227 GTGCATCGTATCTGGCCGACGAGACATCTGTG 2262  
 RESULT 74  
 ABL67952  
 ID ABL67952 standard; DNA: 3370 BP.  
 AC ABL67952;  
 XX 15-MAY-2002 (first entry)  
 DT

XX Ovary cancer related gene sequence SEQ ID NO:6289.  
DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytoskeletal; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
XX gene; ds.  
OS Homo sapiens.  
XX WO200194629-A2.  
PN 13-DEC-2001.  
XX 30-MAY-2001; 2001WO-US10838.  
PF 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX  
XX (AVALON PHARM.  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX WPI; 2002-188264/24.  
XX  
XX Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set  
XX  
XX Claim 1; SEQ ID 6289; 44pp; English.  
XX The present invention describes a method (M1) for screening for an

CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (II) has cytoskeletal  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX  
SQ Sequence 3370 BP; 689 A; 1002 C; 1005 G; 674 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 0.698 Length: 3370  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: Gaps: 0  
  
US-09-397-967-16 (1-1099) x ABL67952 (1-3370)  
  
QY 941 VALHISARGSPLEUALAALARGASMLEUVAL 952  
Db 2353 GTCCACGGAGCTGCTGCCGAAACATCTTGTCG 2388  
  
RESULT 75  
ID ABR49561 standard; DNA; 3415 BP.  
XX ABR49561;  
AC 15-JUL-2002 (first entry)  
DT  
XX Human CDNA 16658 encoding a novel kinase.  
DE  
XX Human; ss; gene; 16658; kinase; cellular proliferative disorder; cancer;  
KW carcinoma; tumour; adenocarcinoma; haematopoietic neoplastic disorder;  
KW leukaemia; lymphoma; brain disorder; cerebral ischaemia; infection;  
KW meningitis; brain abscess; acquire immunodeficiency syndrome; obesity;  
KW AIDS-related myopathy; prion disease; Alzheimer's disease; diabetes;  
KW Parkinson's disease; Huntington's disease; motor neurone disease;  
KW metabolic disorder; anorexia nervosa; pain; inflammation; ischaemia;  
KW irritable bowel syndrome; heart disorder; myocardial infarction;  
KW blood vessel disorder; atherosclerosis; bone metabolism disorder;  
KW osteoporosis; haematopoietic disorder; arthritis.  
XX  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH 23..3415  
FT CDS  
FT /\*tag= a  
FT /product= "kinase 16658"  
FT /note= "This coding sequence is specifically claimed  
FT in claim 1"  
XX  
XX WO200220800-A2.  
XX  
XX 14-MAR-2002.  
XX  
XX 03-AUG-2001; 2001WO-US24601.  
XX  
XX 01-SEP-2000; 2000US-229299P.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX

PI Meyers RA, Sllas-Sanliago I;  
 XX WPI: 2002-351781/38.  
 DR P-PSDB: AAU79456.  
 XX  
 XX  
 PT New human kinase protein for diagnosing and treating disorders  
 PT e.g. cell proliferative, metabolic, cardiovascular, neurological and to  
 PT identify modulators for therapeutic use  
 PS  
 XX Claim 1: Fig 1: 143pp: English.  
 XX  
 CC The invention relates to an isolated human kinase polypeptide  
 CC encoded by the DNAs designated 16558, 14223 and 16002 including  
 CC fragments, homologues and allelic variants. Also included are a host  
 CC cell comprising the DNA, an antibody which selectively binds to the  
 CC novel kinase, a method for producing the novel kinase comprising  
 CC culturing the host cell and recovering the protein, detecting the  
 CC presence of the DNA in a sample comprising contacting the sample  
 CC with a compound which selectively hybridises to the DNA and determining  
 CC whether the compound has bound and identifying compounds which bind to  
 CC and/or modulate the protein comprising contacting the protein with a  
 CC test compound and determining whether the compound has bound to and/or  
 CC modulated the function of the protein (the modulators may be a small  
 CC molecule, a peptide, a phosphopeptide, an antibody or a fragment of the  
 CC full length protein). The nucleic acids, proteins, identified  
 CC modulators and antibodies are useful in the diagnosis, monitoring and  
 CC treatment of a wide range of diseases and disorders (many examples of  
 CC which are listed in the specification) including cellular proliferative  
 CC disorders (e.g. cancers of the lung and breast, carcinomas, tumours,  
 CC adenocarcinomas, haematopoietic neoplastic disorders e.g. leukaemias and  
 CC lymphomas), brain and nerve tissue disorders (e.g. cerebral ischaemia,  
 CC infections such as meningitis, brain abscess, acquired immunodeficiency  
 CC syndrome (AIDS)-related myopathy, prion diseases, Alzheimer's  
 CC disease, Parkinson's disease, Huntington's disease and motor neurone  
 CC disease), metabolic disorders (e.g. obesity, anorexia nervosa and  
 CC diabetes) pain disorders (e.g. associated with infection, inflammation  
 CC ischaemia, irritable bowel syndrome), heart disorders (e.g. myocardial  
 CC infarction), blood vessel disorders (e.g. atherosclerosis), disorders of  
 CC bone metabolism (e.g. osteoporosis) and haematopoietic disorders (e.g.  
 CC arthritis). The present sequence encodes novel human kinase 16558.  
 XX  
 SQ Sequence 3415 BP; 951 A; 777 C; 849 G; 838 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.707 Length: 3415  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x ABK49561 (1-3415)  
 OY 941 VAHHSARGASPLEUALAALAAARGASNIIELEVAL 952  
 |||||||  
 DB 2687 GTTCATCGAGACCTACGCGCTCGGAATATRGTGTC 2722

RESULT 76  
 AAH77997  
 ID AAH77997 standard; DNA: 3427 BP.  
 XX  
 AC AAH77997;  
 XX  
 DT 13-NOV-2001 (first entry)  
 XX  
 DE Nucleotide sequence of human protein kinase SGR251.  
 XX  
 XX Human; protein kinase; cancer; immune disease; cardiovascular disease;  
 KW brain disease; neuronal disease; Alzheimer's disease; chromosome 11;  
 KW Parkinson's disease; multiple sclerosis; metabolic disorder;  
 KW peripheral nervous system disease; amyotrophic lateral sclerosis;  
 KW infection; ocular disease; migraine; pain; sexual dysfunction;  
 KW mood disorder; attention disorder; cognition disorder; hypotension;

KW hypertension; psychotic disorder; dyskinesia; transplant rejection; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200166594-A2.  
 XX  
 PD 13-SEP-2001.  
 XX  
 PF 02-MAR-2001; 2001WO-US06638.  
 XX  
 PR 06-MAR-2000; 2000US-0187150.  
 PR 29-MAR-2000; 2000US-0193404.  
 PR 13-NOV-2000; 2000US-0247013.  
 XX  
 PA (SUGEN) SUGEN INC.  
 XX  
 PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
 PT WPI: 2001-536777/59.  
 XX  
 DR Nucleic acids capable of encoding human polypeptides having a kinase or  
 PT kinase-like activity, useful for diagnosing a disease selected from  
 PT cancers, cardiovascular disease and neuronal-associated diseases (e.g.  
 PT Alzheimer's disease) -  
 XX  
 XX Example 1; Fig 1H-I; 201pp; English.  
 PS  
 XX The present sequence encodes a human protein kinase. The gene is  
 CC located on chromosome 11. The kinase polypeptides are useful for  
 CC diagnosing a disease or disorder selected from cancers (e.g. cancers  
 CC of tissues and cancers of hematopoietic origin), immune-related diseases  
 CC and disorders, cardiovascular disease, brain or neuronal-associated  
 CC diseases (e.g. Alzheimer's disease, Parkinson's disease, multiple  
 CC sclerosis), metabolic disorders, peripheral nervous system diseases,  
 CC amyotrophic lateral sclerosis, viral infections, infections caused by  
 CC prions, infections caused by bacteria, infections caused by fungi,  
 CC ocular diseases, migraines, pain, sexual dysfunction, mood disorders,  
 CC attention disorders, cognition disorders, hypotension, hypertension,  
 CC psychotic disorders, dyskinesias, and organ transplant rejection.  
 CC Kinase inhibitors are useful for treating diseases and disorders  
 CC described above.  
 XX  
 SQ Sequence 3427 BP; 955 A; 797 C; 788 G; 887 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.709 Length: 3427  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x AAH77997 (1-3427)  
 OY 941 VAHHSARGASPLEUALAALAAARGASNIIELEVAL 952  
 |||||||  
 DB 2186 GTTCATCGAGACCTACGCGCTCGGAATATRGTGTC 2221

RESULT 77  
 AAD38862  
 ID AAD38862 standard; CDNA: 3472 BP.  
 XX  
 AC AAD38862;  
 XX  
 DT 23-SEP-2002 (first entry)  
 XX  
 DE Human kinase (PKIN)-19 CDNA.  
 XX  
 XX Human; kinase; cancer; immune system disorder; atherosclerosis;  
 KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;  
 KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;  
 KW development; hepatitis; cardiovascular; hypertension; drug screening;  
 KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;





DB 2189 GTGACCGAGACCTGCTGCCGCAACATCTGTC 2224

RESULT 79

ABK70012

ID ABK70012 standard; DNA: 3583 BP.

XX

AC ABK70012;

XX

DT 15-JUL-2002 (first entry)

XX

DE cDNA encoding human Pro peptide #52.

XX

KH Human; ss: gene; PRO; secreted protein; transmembrane protein;

KW genetic disorder; tumour; cancer.

XX

OS Homo sapiens.

XX

PN WO20024888-A2.

XX

PD 28-MAR-2002.

XX

PE 29-AUG-2001; 2001WO-US27099.

XX

PR 01-SEP-2000; 2000US-229896P.

PR 05-SEP-2000; 2000US-230621P.

PR 22-SEP-2000; 2000US-235147P.

PR 10-NOV-2000; 2000WO-US30873.

PR 12-JAN-2001; 2001US-261878P.

PR 16-JAN-2001; 2001US-261910P.

PR 16-JAN-2001; 2001US-261939P.

PR 25-JAN-2001; 2001US-262150P.

PR 02-FEB-2001; 2001US-264395P.

PR 09-FEB-2001; 2001US-266421P.

PR 28-FEB-2001; 2001US-267623P.

PR 09-MAR-2001; 2001US-274399P.

PR 03-APR-2001; 2001US-280982P.

PR 04-APR-2001; 2001US-282192P.

PR 09-MAY-2001; 2001US-282199P.

PR 25-MAY-2001; 2001US-290589P.

PR 01-JUN-2001; 2001WO-US17092.

PR 20-JUN-2001; 2001WO-US19692.

PR 29-JUN-2001; 2001WO-US21066.

PR 09-JUL-2001; 2001WO-US21735.

XX

PA (GETH ) GENENTECH INC.

XX

PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;

PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;

PI Fong S;

XX

DR WPI: 2002-362426/39.

DR P-PSDB: ABG34081.

XX

PT New PRO polypeptides and polynucleotides encoding the polypeptides,

PT useful in gene therapy, chromosome identification, tissue typing, or

PT for genetic analysis of individuals with genetic disorders

XX

PS Claim 2; Figure 103; 218pp; English.

CC This invention relates to the cDNA and protein sequences of novel

CC secreted and transmembrane polypeptides PRO polypeptides. The

CC invention also comprises a method for producing the proteins of the

CC invention by recombinant means and antibodies specific for the protein

CC proteins of the invention. The antibody may be used for detecting the PRO

CC polynucleotides may be used as hybridisation probes for a cDNA library

CC to isolate the full-length PRO cDNA or to isolate other cDNAs, to

CC construct hybridisation probes for mapping the gene which encodes that

CC PRO and for genetic analysis of individuals with genetic disorders. In

CC assays to identify other proteins or molecules involved in binding

CC reaction, to generate transgenic animals or knock-out animals which in

CC turn are useful in the development and screening of therapeutically

CC useful reagents, for chromosome identification, and tissue typing. The

CC PRO polypeptides are useful in gene therapy, and as molecular weight

CC markers for protein electrophoresis purposes. The sequences may

CC also be used to detect overexpression on PRO polypeptides in cancerous

CC tumours and for screening for differentially expressed genes using

CC microarray technology. The present sequence represents a cDNA encoding

CC a human PRO protein of the invention.

XX

SQ Sequence 3583 BP; 1019 A; 816 C; 853 G; 895 T; 0 other:

Alignment Scores:

Pred. No.:	0.74	Length:	3583
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	24	Gaps:	0

US-09-397-967-16 (1-1099) x ABK70012 (1-3583)

OY 941 ValHisArgAspLeuAlaIaArgAsnIleLeuVal 952

DB 2633 GTTCATCGAGACCTAGCGCTCGAATATACCTGCTC 2668

RESULT 80

AAQ90658

ID AAQ90658 standard; cDNA: 3591 BP.

XX

AC AAQ90658;

XX

DT 11-NOV-1995 (first entry)

XX

DE Eph-related PTK Cck10+ cDNA.

XX

KW Cck10+; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;

KW prognosis; ss.

XX

OS Gallus sp.

XX

FT Key Location/Qualifiers

FT CDS 2..2968

FT /\*tag= a

XX

PN WO9515375-A.

XX

PD 08-JUN-1995.

XX

PF 07-SEP-1994; 94WO-US10140.

XX

PR 03-DEC-1993; 93US-0162809.

XX

PA (LJOL-) LA JOLLA CANCER RES FOUND.

XX

PI Pasquale EB, Sajjadi FG;

XX

DR WPI: 1995-215256/28.

DR P-PSDB: AAR75710.

XX

PT Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing

PT cancer.

XX

PS Claim 2; Page 78-82; 129pp; English.

XX

CC Novel Eph-related PTK cDNA clone Cck10 (given in AAQ90656) and a

CC variant clone, Cck10+ (AAQ90658), whose product contains a 15-amino

CC acid insertion in the juxtamembrane domain, were isolated from a

CC chick embryo library in lambda gt11. Cck10 and Cck10+ may

CC originate from the same gene by alternative splicing. Cck10

CC expression was prominent in the kidney, and to a lesser extent in

CC the lung.

XX

SQ Sequence 3591 BP; 753 A; 1112 C; 1042 G; 684 T; 0 other;

10

DD.

Query Match:  
DB:

Query Match:  
DB:

```

US-09-397-967-16 (1-1099) x AAT72320 (1-3663)
QY 941 VALHISARGSPLEUALAALARGASNIIEUVAL 952
   |||||||
DB 2593 GTCCATAGAGACCTGGCTGCCAGAACATCTTGTC 2628
RESULT 83
AA062461
ID AA062461 standard; cDNA; 3751 BP.
XX
AC AA062461;
XX
DT 09-NOV-1994 (first entry)
XX
DE Human embryonal kinase 2 receptor.
XX
KM Embryonal kinase; HEK; protein tyrosine kinase; PTK; tumour;
KM cancer; therapy; amplification; primer; polymerase chain reaction;
KM PCR; ss.
XX
OS Homo sapiens.
XX
FH Key 1.2973
FT CDS /tag= a
XX
PD DE4233782-A.
XX
PD 14-APR-1994.
XX
PF 07-OCT-1992; 92DE-4233782.
XX
PR 07-OCT-1992; 92DE-4233782.
XX
PA (CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.
XX
PI Holtrich U, Ruebsamen-Waigmann H, Strebhardt K;
XX
DR WPI: 1994-127194/16.
DR P-PSDB: AAR51899.
XX
PT Human embryonal kinase 2-receptor protein - useful in tumour
PT diagnosis and therapy
XX
PS Claim 4; Page 7-10; 11pp; German.
XX
CC RNA from human embryonic tissue was isolated. With the use of
CC primer P6(4) PTK-specific cDNA was synthesised. The cDNA was
CC amplified using primers P6(4) and N5. A 2097 bp DNA fragment was
CC obtained. Primers E3, P12 and E6 were then used in the isolation of
CC the C-terminal of the HEK2 receptor gene.
XX
SQ Sequence 3751 BP; 739 A; 1150 C; 1123 G; 739 T; 0 other;

Alignment Scores:
Pred. No.: 0.772 Length: 3751
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 15 Gaps: 0

US-09-397-967-16 (1-1099) x AA062461 (1-3751)
QY 941 VALHISARGSPLEUALAALARGASNIIEUVAL 952
   |||||||
DB 2239 GTGCACCGGACCTGGCTGCCAGAACATCTTGTC 2274
RESULT 84
AAA88548
ID AAA88548 standard; cDNA; 3768 BP.
XX
AC AAA88548;

```

```

XX
DT 22-JAN-2001 (first entry)
XX
DE Human CASB616 cDNA.
XX
KM CASB616; EPHB2; ERK; EPH3; EPH3T3; DRT; HEK5; EPHB2V;
KM receptor protein tyrosine kinase; human; antigen; colon cancer;
KM ovary cancer; tumour; autoimmune disease; vaccine; gene therapy;
KM diagnosis; ss.
XX
OS Homo sapiens.
XX
FH Key 105..3068
FT CDS /tag= a
FT CDS /transl_except= (pos:2973..2975,aa:Leu)
XX
PN WO200053216-A2.
XX
PD 14-SEP-2000.
XX
PF 28-FEB-2000; 2000WO-BP01587.
XX
PR 05-MAR-1999; 99GB-0005124.
XX
PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Vinals De Bassols XC;
XX
DR WPI: 2000-587384/55.
DR P-PSDB: AAB19590.
XX
PT Vaccine composition for treating ovarian and colon cancer, comprises
PT CASB616 polypeptides, polynucleotides or antigen presenting cells
PT expressing the polypeptides
XX
PS Claim 3; Page 40-41; 57pp; English.
XX
CC The present sequence is that of cDNA coding for human CASB616 (see
CC AAB19590), a member of the EPH and EPH-related family of receptor
CC protein tyrosine kinases. CASB616 is also known as EPHB2, ERK,
CC EPH3, EPH3T3, DRT, HEK5 and EPHB2V. CASB616 polypeptides and
CC polynucleotides are important immunogens for specific prophylactic
CC or therapeutic immunization against tumours, especially colon
CC cancer (claimed) and ovarian cancer. They are specifically
CC expressed or highly over-expressed in tumours compared to normal
CC cells and can thus be targeted by antigen-specific immune
CC mechanisms leading to destruction of the tumour cells. They can
CC also be used to diagnose the occurrence of tumour cells. Their
CC inappropriate expression can also cause an induction of autoimmune
CC responses, which can be corrected through vaccination using the
CC CASB616 polypeptides or polynucleotides.
XX
SQ Sequence 3768 BP; 872 A; 1129 C; 1058 G; 709 T; 0 other;

Alignment Scores:
Pred. No.: 0.775 Length: 3768
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 21 Gaps: 0

US-09-397-967-16 (1-1099) x AAA88548 (1-3768)
QY 941 VALHISARGSPLEUALAALARGASNIIEUVAL 952
   |||||||
DB 2334 GTTCACCGTGTGCTGCCGCGAACATCTTGTC 2369
RESULT 85
AA090655
ID AA090655 standard; cDNA; 3776 BP.
XX

```

AC AA090655;  
 XX 11-NOV-1995 (first entry)  
 DT  
 XX Eph-related PTK Cdk9 CDNA.  
 DE  
 XX Cdk9; Eph: protein tyrosine-kinase; PTK; cancer; diagnosis;  
 KW prognosis; ss.  
 KM  
 XX Gallus sp.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 290..311  
 FT /\*tag= a  
 FT  
 XX  
 XX MO9515375-A.  
 PN  
 XX  
 XX PD 08-JUN-1995.  
 XX  
 XX PF 07-SEP-1994; 94MO-US10140.  
 XX  
 XX PR 03-DEC-1993; 93US-0162809.  
 XX  
 XX (LJOL-) LA JOLLA CANCER RES FOUND.  
 PA  
 XX Pasquale EB, Sajjadi FG;  
 PI  
 XX WPI; 1995-215256/28.  
 DR  
 XX P-PSDB; AAR75707.  
 XX  
 XX Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing  
 PT cancer.  
 PT  
 XX  
 XX Claim 2; Page 50-53; 129pp; English.  
 PS  
 XX Novel Eph-related PTK cDNA clone Cdk9 (given in AA090655) was  
 CC isolated from a chick embryo library in lambda. Cdk9 protein  
 CC (AAR75707) is closely related to Cdk5 (AAR75712). In adult tissues,  
 CC Cdk9 expression is predominant in the thymus and detectable in  
 CC brain, retina, kidney, lung and heart.  
 CC  
 XX  
 SQ Sequence 3776 BP; 1005 A; 881 C; 941 G; 949 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.777 Length: 3776  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 16 Gaps: 0  
 US-09-397-967-16 (1-1099) x AA090655 (1-3776)  
 OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 DB 2477 GTGCATCGTATCTCGCAGCAGTACATCTTACTC 2512  
 RESULT 86  
 ID ABA09147 standard: cDNA; 3836 BP.  
 XX  
 AC ABA09147;  
 XX  
 DT 11-JAN-2002 (first entry)  
 DE Human protein Tyr kinase receptor homologue cDNA, SEQ ID NO:923.  
 XX  
 XX Human; cytokine; cell proliferation; cell differentiation; growth factor;  
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 KW chronic inflammatory condition; proliferative retinopathy;

KW atherosclerosis; coronary heart disease; arterial ischaemia;  
 KW bone disorder; osteoporosis; vascular growth disorder;  
 KW tissue regeneration; wound healing; infection; immune disorder;  
 KW cell culture; drug screening; gene therapy; antiinflammatory;  
 KW antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;  
 KW cytostatic; osteopathic; vasotropic; cardiatic; virucide; antibacterial;  
 KW antifungal; vulnery; antitumor; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX MO200157188-A2.  
 PN  
 XX  
 XX PD 09-AUG-2001.  
 XX  
 XX PF 05-FEB-2001; 2001MO-US03800.  
 XX  
 XX PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX  
 XX Tang YT, Liu C, Drmanac RT;  
 PI  
 XX WPI; 2001-457740/49.  
 DR  
 XX P-PSDB; ABB11903.  
 DR  
 XX Human proteins and DNA encoding sequences useful for preventing,  
 PT treating or ameliorating a medical condition in a mammalian subject  
 PT e.g. arthritis and cancer -  
 XX  
 XX Claim 1; Page 798-799; 1963pp; English.  
 PS  
 XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
 CC sequences ABA08325-ABA09574 represent nucleic acids encoding them. The  
 CC invention also relates to vectors and recombinant host cells comprising a  
 CC nucleotide of the invention, methods of producing the novel polypeptides,  
 CC antibodies against the polypeptides, methods of detecting the nucleotides  
 CC or polypeptides in a sample, and methods of identifying compounds which  
 CC bind to polypeptides of the invention. Although novel, many of the  
 CC polypeptides of the invention have homology to known proteins, thereby  
 CC giving an insight into their probable biological activities, and hence  
 CC potential therapeutic applications. The polypeptides of the invention may  
 CC have various activities, including cytokine, cell proliferation or cell  
 CC differentiation activities; stem cell growth factor activity;  
 CC haematopoiesis regulatory activity; tissue growth activity;  
 CC immunomodulatory activity; activin- or inhibin-related activities;  
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
 CC thrombolytic activities; receptor or ligand activities; or may be  
 CC involved in oncogenesis, cancer cell proliferation or metastasis.  
 CC Depending on their biological activities, polypeptides and nucleotides of  
 CC the invention are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g., by protein or gene therapy. Such conditions include  
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
 CC vascular growth. Polypeptides involved with tissue regeneration and  
 CC repair (or nucleic acids encoding them) may be used to promote wound  
 CC healing (e.g., of burns, incisions and ulcers), while those with  
 CC immunomodulatory activities may be used in the treatment of viral,  
 CC bacterial and fungal infections in addition to immune disorders.  
 CC Polypeptides with growth factor activity may be used in cell cultures to  
 CC promote cell growth. For example, such polypeptides may be used to  
 CC manipulate stem cells in culture to give rise to neuroepithelial cells  
 CC that can be used to augment or replace cells damaged by illness,  
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
 CC may also be used in the diagnosis of the above conditions, and in drug  
 CC screening techniques. The present sequence represents a cDNA encoding a  
 CC novel human polypeptide of the invention.  
 XX  
 SQ Sequence 3836 BP; 759 A; 1190 C; 1140 G; 746 T; 1 other;  
 Alignment Scores:

Pred. No.: 0.788 Length: 3836  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservatve: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 22 Gaps: 0

US-09-397-967-16 (1-1099) x ABA09147 (1-3836)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952  
Db 2290 GTGCACCGACCTGCTGCAACATCTTGTTC 2325

RESULT 87  
AAD27896  
ID AAD27896 standard; CDNA: 3945 BP.

XX AAD27896;  
XX  
XX  
XX 31-MAY-2002 (first entry)  
XX  
XX Human EphB4 CDNA.  
XX  
XX Artery; vein; endothelial cell; obstructed blood vessel; coronary artery;  
XX atherosclerosis; graft; gene therapy; human; EphB4; arterial marker; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200211785-A2.  
XX  
XX 14-FEB-2002.  
XX  
XX 03-AUG-2001; 2001WO-US24405.  
XX  
XX 03-AUG-2000; 2000US-222759P.  
XX  
XX (UTAH ) UNIV UTAH.  
XX  
XX L1 DY:  
XX  
XX WPI, 2002-217155/27.  
XX  
XX  
XX Inducing arterial morphology in vein by contacting endothelial cells in  
XX vein to a polynucleotide encoding gene capable of inducing endothelial  
XX remodelling to transfer the polynucleotide into the cells  
XX  
XX  
XX Claim 6; Page 30-34; 39pp; English.

CC The invention relates to a method of inducing arterial morphology in a  
CC vein, comprising contacting endothelial cells in the vein to at least  
CC one polynucleotide encoding a gene that is capable of inducing  
CC endothelial remodelling for a time sufficient to transfer the  
CC polynucleotide into the endothelial cells. A method for treating a  
CC patient having an obstructed blood vessel, e.g. coronary artery  
CC affected by atherosclerosis, is also provided which comprises providing  
CC a graft comprising endothelial cells, contacting the endothelial cells of  
CC the graft to the polynucleotide, removing a section of the obstructed  
CC blood vessel, and grafting the graft in place of the removed section of  
CC the obstructed blood vessel, where a graft is provided by harvesting a  
CC section of a vein from the patient. The present sequence is human  
CC EphB4 CDNA used in the invention for endothelial remodelling.  
CC The EphB4 gene encodes an arterial specific molecular marker.  
XX  
XX Sequence 3945 BP; 766 A; 1186 C; 1212 G; 781 T; 0 other;

Alignment Scores:  
Pred. No.: 0.809 Length: 3945  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservatve: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 24 Gaps: 0

US-09-397-967-16 (1-1099) x AAD27896 (1-3945)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952  
Db 2294 GTGCACCGACCTGCTGCAACATCTTGTTC 2329

RESULT 88  
ABL68537  
ID ABL68537 standard; DNA: 3945 BP.

XX ABL68537;  
XX  
XX  
XX 15-MAY-2002 (first entry)  
XX  
XX Kidney cancer related gene sequence SEQ ID NO:6874.  
XX  
XX  
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
XX cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
XX gene; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200194629-A2.  
XX  
XX 13-DEC-2001.  
XX  
XX 30-MAY-2001; 2001WO-US10838.  
XX  
XX 05-JUN-2000; 2000US-209473P.  
XX  
XX 05-JUN-2000; 2000US-209531P.  
XX  
XX 18-SEP-2000; 2000US-233133P.  
XX  
XX 18-SEP-2000; 2000US-233617P.  
XX  
XX 20-SEP-2000; 2000US-234009P.  
XX  
XX 20-SEP-2000; 2000US-234034P.  
XX  
XX 20-SEP-2000; 2000US-234052P.  
XX  
XX 22-SEP-2000; 2000US-234059P.  
XX  
XX 22-SEP-2000; 2000US-234567P.  
XX  
XX 25-SEP-2000; 2000US-234923P.  
XX  
XX 25-SEP-2000; 2000US-234924P.  
XX  
XX 25-SEP-2000; 2000US-235077P.  
XX  
XX 25-SEP-2000; 2000US-235082P.  
XX  
XX 25-SEP-2000; 2000US-235134P.  
XX  
XX 25-SEP-2000; 2000US-235280P.  
XX  
XX 26-SEP-2000; 2000US-235637P.  
XX  
XX 26-SEP-2000; 2000US-235638P.  
XX  
XX 27-SEP-2000; 2000US-235711P.  
XX  
XX 27-SEP-2000; 2000US-235720P.  
XX  
XX 27-SEP-2000; 2000US-235840P.  
XX  
XX 27-SEP-2000; 2000US-235863P.  
XX  
XX 28-SEP-2000; 2000US-236028P.  
XX  
XX 28-SEP-2000; 2000US-236032P.  
XX  
XX 28-SEP-2000; 2000US-236033P.  
XX  
XX 28-SEP-2000; 2000US-236034P.  
XX  
XX 28-SEP-2000; 2000US-236109P.  
XX  
XX 28-SEP-2000; 2000US-236111P.  
XX  
XX 29-SEP-2000; 2000US-236842P.  
XX  
XX 29-SEP-2000; 2000US-236891P.  
XX  
XX 02-OCT-2000; 2000US-237172P.  
XX  
XX 02-OCT-2000; 2000US-237173P.  
XX  
XX 02-OCT-2000; 2000US-237278P.  
XX  
XX 02-OCT-2000; 2000US-237294P.  
XX  
XX 02-OCT-2000; 2000US-237295P.  
XX  
XX 02-OCT-2000; 2000US-237316P.  
XX  
XX 03-OCT-2000; 2000US-237425P.  
XX  
XX 03-OCT-2000; 2000US-237598P.  
XX  
XX 03-OCT-2000; 2000US-237604P.  
XX  
XX 03-OCT-2000; 2000US-237606P.  
XX  
XX 03-OCT-2000; 2000US-237608P.  
XX  
XX 01-NOV-2000; 2000US-244687P.  
XX  
XX 01-NOV-2000; 2000US-245084P.  
XX  
XX (AVAL-) AVALON PHARM.

```

XX  Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI  Soppet DR, Weaver Z;
XX  WPI: 2002-188264/24.
XX
PT  Screening for anti-neoplastic agent involves exposing cells to a
PT  chemical agent to be tested for anti-neoplastic activity, and
PT  determining a change in expression of a gene of a signature gene set
XX
PS  Claim 1; SEQ ID 6874; 44pp; English.
XX
CC  The present invention describes a method (M1) for screening for an
CC  anti-neoplastic agent. The method involves exposing cells to a chemical
CC  agent to be tested for anti-neoplastic activity, determining a change in
CC  expression of at least one gene (I) of a signature gene set, where (i)
CC  comprises a sequence (S) selected from 8447 sequences (given in ABU61664
CC  to ABU70110), or is at least 95% identical to (S), where a change in
CC  expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC  activity and can be used in gene therapy. M1 can be used for screening
CC  an anti-neoplastic agent, and can be used for producing a product which
CC  is the data collected with respect to the anti-neoplastic agent as a
CC  result of M1, and the data is sufficient to convey the chemical
CC  structure and/or properties of the agent. M1 can be used in the
CC  treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC  oesophagagal, ovarian, kidney, prostate or pancreatic cancer,
CC  adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC  infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC  carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ  Sequence 3945 BP; 763 A; 1188 C; 1215 G; 779 T; 0 other;

Alignment Scores:
Pred. No.:      0.809      Length:      3945
Score:          12.00      Matches:      12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:    1.09%      Indels:      0
DB:             24         Gaps:        0

US-09-397-967-16 (1-1099) x ABU68537 (1-3945)

QY  941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
    |||||||
DB  2294 GTCCACCGAGACGTGGCTGCTGCACATCCTAGTC 2329

RESULT 89
ID  AAA88549 standard; cDNA: 3949 BP.
XX
AC  AAA88549;
XX
DT  22-JAN-2001 (first entry)
XX
DE  Human CASB616 cDNA.
XX
KW  CASB616; EPHB2; ERK; EPH3; EPH3; DRT; HEK5; EPHB2V;
KW  receptor protein tyrosine kinase; human; antigen; colon cancer;
KW  ovary cancer; tumour; autoimmune disease; vaccine; gene therapy;
XX
OS  Homo sapiens.
XX
FH  Key Location/Qualifiers
FT  CDS 26..3193
FT  FT /*tag= a
XX
XX  WO200053216-A2.
XX  14-SEP-2000.
XX
PF  28-FEB-2000; 2000WO-EP01587.
XX

```

```

PR  05-MAR-1999; 99GB-0005124.
XX
PA  (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI  Vinals De Bassols YC;
XX
DR  WPI: 2000-587384/55.
DR  P-PSDB; AAB19591.
XX
PT  Vaccine composition for treating ovarian and colon cancer, comprises
PT  CASB616 polypeptides, polynucleotides or antigen presenting cells
PT  expressing the polypeptides
XX
PS  Claim 3; Page 41-42; 57pp; English.
XX
CC  The present sequence is that of cDNA coding for human CASB616 (see
CC  AAB19591), a member of the EPH and EPH-related family of receptor
CC  protein tyrosine kinases. CASB616 is also known as EPHB2, ERK,
CC  EPH3, EPH3, DRT, HEK5 and EPHB2V. CASB616 polypeptides and
CC  polynucleotides are important immunogens for specific prophylactic
CC  or therapeutic immunization against tumours, especially colon
CC  cancer (claimed) and ovarian cancer. They are specifically
CC  expressed or highly over-expressed in tumours compared to normal
CC  cells and can thus be targeted by antigen-specific immune
CC  mechanisms leading to destruction of the tumour cells. They can
CC  also be used to diagnose the occurrence of tumour cells. Their
CC  inappropriate expression can also cause an induction of autoimmune
CC  responses, which can be corrected through vaccination using the
CC  CASB616 polypeptides or polynucleotides.
XX
SQ  Sequence 3949 BP; 934 A; 1142 C; 1105 G; 768 T; 0 other;

Alignment Scores:
Pred. No.:      0.81      Length:      3949
Score:          12.00      Matches:      12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:    1.09%      Indels:      0
DB:             21         Gaps:        0

US-09-397-967-16 (1-1099) x AAA88549 (1-3949)

QY  941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
    |||||||
DB  2252 GTTCACCGAGACGTGGCTGCCGCCAATCCTCGTC 2287

RESULT 90
ID  AAQ49757 standard; DNA: 3969 BP.
XX
AC  AAQ49757;
XX
DT  10-MAR-1994 (first entry)
XX
DE  PTK gene HPTK5.
XX
KW  PTK; protein tyrosine kinase; catalytic domain; c-kit; hepatoma cell;
KW  amplification; primer; polymerase chain reaction; PCR; ds.
XX
OS  Homo sapiens.
XX
FH  Key Location/Qualifiers
FT  CDS 89..3969
FT  FT /*tag= a
FT  FT /note= "the first stop codon is at position 3051-3053"
FT  FT misc_RNA 701..800
FT  FT /*tag= b
FT  FT /note= "100 bases are missing from the sequence
FT  FT of the specification and entered
FT  FT as (N)100 by the indexer"
XX
XX  WO9315201-A.
XX

```

PD 05-AUG-1993.  
 XX  
 PF 22-JAN-1993; 93WO-US00586.  
 XX  
 PR 22-JAN-1992; 92US-0826935.  
 XX  
 PA (NEW-) NEW ENGLAND DEACONESS HOSPITAL.  
 XX  
 PI Avraham H, Cowley S, Groopman J, Scadden D;  
 XX  
 DR WPI; 1993-320330/40.  
 DR P-PSDB; AAR41896.  
 XX  
 PT New protein tyrosine kinase genes and proteins encoded by genes -  
 PT are of human mega-karyocytic origin  
 XX  
 PS Claim 2; Fig 8; 60pp; English.  
 XX  
 CC PTK genes were identified using two sets of degenerative  
 CC oligonucleotide primers: a first set which amplifies all PTK DNA  
 CC segments (AAQ9743-44), and a second set which amplifies highly  
 CC conserved sequences present in the catalytic domain of the c-kit  
 CC subgroup of PTKs (AAQ49745-46). The PTK genes identified are described  
 CC in AAQ49747-57 and AAR41897-02.  
 CC The HPTK5 gene is expressed in human hepatoma cells.  
 XX  
 SQ Sequence 3969 BP; 777 A; 1149 C; 1180 G; 763 T; 100 other;  
 Alignment Scores:  
 Pred. No.: 0.814 Length: 3969  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 14 Gaps: 0  
 US-09-397-967-16 (1-1099) x AAQ49757 (1-3969)  
 QY 941 VALHISARGSPLEUALAALARGASNIILEUVAL 952  
 DB 2298 GTCCACCGAGACCTGCTGCTGCACACATCTTAGTC 2333  
 RESULT 91  
 AAT03099  
 ID AAT03099 standard; DNA; 3969 BP.  
 AC AAT03099;  
 XX  
 DT 14-FEB-1996 (first entry)  
 XX  
 DE Protein tyrosine-kinase HPTK5 gene.  
 XX  
 KW Protein tyrosine-kinase; PTK; HPTK5; hepatoma transmembrane kinase;  
 KW agonist; cell growth; differentiation; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 90..3053  
 FT sig\_peptide 90..134  
 FT mat\_peptide 135..3050  
 FT /\*tag= b  
 FT /\*tag= c  
 FT /\*product= HPTK5  
 XX  
 PN WO9527061-A1.  
 XX  
 PD 12-OCT-1995.  
 XX  
 PF 04-APR-1995; 95WO-US04228.  
 XX  
 PR 04-APR-1994; 94US-0222616.  
 PR

XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;  
 PI Wood WI;  
 XX  
 DR WPI; 1995-366160/47.  
 DR P-PSDB; AAR85930.  
 XX  
 PT Agonist antibodies which activate specific protein tyrosine  
 PT kinase(s) - also activate chimeric proteins of kinase extracellular  
 PT domain and Ig constant domain, useful for studying, and therapeutic  
 PT modulation of, cell growth and differentiation  
 XX  
 PS Disclosure; Page 62-67; 125pp; English.  
 XX  
 CC DNA probes based on protein tyrosine-kinase (PTK) sequences were used  
 CC to screen cDNA libraries to identify novel PTK genes. The HPTK5  
 CC gene (AAT03099) was expressed in human hepatoma cells and encoded a  
 CC protein (AAR85930) showing homology to known PTKs. The gene was used to  
 CC construct an HPTK5 extracellular domain-IgG FC fusion protein.  
 XX  
 SQ Sequence 3969 BP; 786 A; 1187 C; 1214 G; 782 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.814 Length: 3969  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.09% Indels: 0  
 DB: 16 Gaps: 0  
 US-09-397-967-16 (1-1099) x AAT03099 (1-3969)  
 QY 941 VALHISARGSPLEUALAALARGASNIILEUVAL 952  
 DB 2298 GTCCACCGAGACCTGCTGCTGCACACATCTTAGTC 2333  
 RESULT 92  
 AAO90972  
 ID AAO90972 standard; cDNA to mRNA; 4027 BP,  
 AC AAO90972;  
 XX  
 DT 24-NOV-1995 (first entry)  
 XX  
 DE Protein p140 cDNA from rat skeletal muscle myoblast cell line L6.  
 XX  
 KW Protein p140; insulin; tyrosine phosphorylation; ss.  
 XX  
 OS Rattus rattus.  
 XX  
 PN EP659883-A.  
 XX  
 PD 28-JUN-1995.  
 XX  
 FE 24-NOV-1994; 94EP-0118524.  
 XX  
 PR 24-NOV-1993; 93JP-0315806.  
 XX  
 PA (ONOV ) ONO PHARM CO LTD.  
 XX  
 PI Kitagawa K, Ohno H, Tajima H;  
 XX  
 DR WPI; 1995-226291/30.  
 XX  
 PT Isolated protein p140 polypeptide - and treatment of diabetes based  
 PT on tyrosine phosphorylation of protein p140.  
 XX  
 PS Claim 5; Page 25-27; 42pp; English.  
 CC  
 CC A cDNA library was established from rat skeletal myoblast cell line  
 CC L6. DNA fragments of approx. 400 bp were recovered and subjected to

CC cloning. 20 plasmids were sequenced. Sequence data of cDNA fragments  
 CC were constructed to the linkage sequences with the DNA programme  
 CC DNASTIS. The basic sequence protrayed in AAQ90972 was hence constructed.  
 CC From sequence data of the whole cDNA length, the ORF was determined.  
 CC The AA sequence was further translated and the sequence thus  
 CC established is illustrated in AAR75843. One of the frames possesses  
 CC the 2993-bp ORF, that was approximated to 3000 bp of the whole ORF  
 CC length of the ECK family.

XX Sequence 4027 BP; 885 A; 1141 C; 1166 G; 835 T; 0 other;  
 SQ

## Alignment Scores:

Pred. No.:	0.825	Length:	4027
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAQ90972 (1-4027)

QY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952  
 |||||||||||||||||||||||||||||||||||  
 DB 2509 GTGCACCGTGACCTGCTGCCGCAACATCTTGTG 2544

## RESULT 93

AAQ90982  
 ID AAQ90982 standard; cDNA to mRNA; 4027 BP.

XX AAQ90982;

XX 24-NOV-1995 (first entry)

XX Protein p140 cDNA from rat skeletal muscle myoblast cell line L6.

XX Protein p140; insulin; tyrosine phosphorylation; ss.

XX Rattus rattus.

XX Key Location/Qualifiers  
 FH CDS 262..3243  
 FT /\*tag= a

XX EP659883-A.

XX 28-JUN-1995.

XX 24-NOV-1994; 94EP-0118524.

XX 24-NOV-1993; 93JP-0315806.

XX (ONOV) ONO PHARM CO LTD.

XX Kitagawa K, Ohno H, Tajima H;

XX WPI: 1995-226291/30.

XX P-PSDB; AAR75844.

XX Isolated protein p140 polypeptide - and treatment of diabetes based  
 PT on tyrosine phosphorylation of protein p140.

XX Example; Page 28-33; 42pp; English.

XX AAQ90982 does not appear to be referred to in the spec. although  
 CC it is listed in the sequence listings (SEQ ID No. 4). It is the  
 CC same length as SEQ ID No. 3 and may be identical to it (see  
 CC AAQ90972).

XX Sequence 4027 BP; 885 A; 1141 C; 1166 G; 835 T; 0 other;  
 SQ

## Alignment Scores:

Pred. No.:	0.825	Length:	4027
Score:	12.00	Matches:	12

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAQ90982 (1-4027)

QY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952  
 |||||||||||||||||||||||||||||||||||  
 DB 2509 GTGCACCGTGACCTGCTGCCGCAACATCTTGTG 2544

## RESULT 94

AAQ90660  
 ID AAQ90660 standard; DNA; 4049 BP.

XX AAQ90660;

XX 11-NOV-1995 (first entry)

XX Eph-related PTK Cek5.

XX Cek5; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;

XX prognosis; ss.

XX Gallus sp.

XX Key Location/Qualifiers  
 FH CDS 10..2997  
 FT /\*tag= a

XX WO9515375-A.

XX 08-JUN-1995.

XX 07-SEP-1994; 94WO-0510140.

XX 03-DEC-1993; 93US-0162809.

XX (LJOL-) LA JOLLA CANCER RES FOUND.

XX Pasquale EB, Sajjadi FG;

XX WPI: 1995-215256/28.

XX P-PSDB; AAR75712.

XX Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing  
 PT cancer.

XX Disclosure; Page 92-96; 129pp; English.

XX Probes derived from the Eph-related PTKs Cek4 (AAQ90659) and Cek5  
 CC (AAQ90660) were used to isolate novel cDNA clones (AAQ90652-58,  
 CC AAQ90661-62) from chicken embryo and embryonic brain libraries.

XX Sequence 4049 BP; 1010 A; 1011 C; 1072 G; 956 T; 0 other;  
 SQ

## Alignment Scores:

Pred. No.:	0.829	Length:	4049
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	16	Gaps:	0

US-09-397-967-16 (1-1099) x AAQ90660 (1-4049)

QY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952  
 |||||||||||||||||||||||||||||||||||

DB 2263 GTGCACCGTGACCTGCTGCCGCAACATCTTGTG 2298

## RESULT 95

AAQ90657  
 ID AAQ90657 standard; DNA; 4097 BP.



```

XX AC AA090657;
XX
XX 11-NOV-1995 (first entry)
XX DE Eph-related PTK Cdk5+.
XX CEK5+; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
XX prognosis; ss.
XX OS Gallus sp.
XX
XX Key Location/Qualifiers
XX FT 10..3045
XX FT /*tag= a
XX
XX MO9515375-A.
XX
XX 08-JUN-1995.
XX
XX 07-SEP-1994; 94WO-US10140.
XX
XX 03-DEC-1993; 93US-0162809.
XX
XX (LJOL-) LA JOLLA CANCER RES FOUND.
XX
XX Pasquale EB, Sajjadi FG;
XX
XX WPI: 1995-215256/28.
XX P-PSDB: AAR75709.
XX
XX EPH-related protein tyrosine kinase(s) - for monitoring and diagnosing
XX cancer.
XX
XX Claim 2; Page 71-75; 129pp; English.
XX
XX A cDNA clone encoding a novel variant of EPH-related PTK Cdk5,
XX CEK5+ (AA090657), was isolated from a chick embryo library in
XX CC lambda gt11. CEK5+ protein (AAR75709) contains a 16-amino acid
XX CC insertion in the juxtamembrane domain, and be a result of
XX CC alternative splicing. CEK5+ is exclusively expressed in the CNS.
XX
XX SO Sequence 4097 BP; 1014 A; 1029 C; 1087 G; 967 T; 0 other;

Alignment Scores:
Pred. No.: 0.839 Length: 4097
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 16 Gaps: 0

US-09-397-967-16 (1-1099) x AA090657 (1-4097)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2311 GTGCACCGGGACCTGCTGCCGCAACATCTGTGTC 2346

RESULT 96
AA053471
ID AA053471 standard; CDNA: 4281 BP.
XX
XX AC AA053471;
XX
XX 16-JUN-1994 (first entry)
XX
XX elk CDNA.
XX
XX Lambda gt11; expression vector; lambda-BI-Elk; protein tyrosine kinase;
XX KW Elk; B1; Eph; subfamily; receptor-like tyrosine kinase; eph; eck;
XX phosphorylation; phosphorylated kinase insert domain; growth factor;
XX receptor kinase; platelet-derived growth factor receptor; ss.
XX

```

```

OS Rattus rattus.
XX
XX Key Location/Qualifiers
XX FT CDS 367..3321
XX FT /*tag= a
XX FT /product= Elk
XX
XX CA2083521-A.
XX
XX 01-OCT-1993.
XX
XX 23-NOV-1992; 92CA-2083521.
XX
XX 31-MAR-1992; 92US-0861390.
XX
XX (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX
XX Letwin K, Pawson A, Reedijk M;
XX
XX WPI: 1993-406300/51.
XX P-PSDB: AAR44513.
XX
XX Expression of phosphorylated exogenous protein - in host cells
XX transformed with two vectors, one for the protein, the other for
XX catalytic domain of protein kinase
XX
XX Disclosure: Fig 3; 55pp; English.
XX
XX This sequence represents the elk cDNA which encodes the protein
XX CC tyrosine kinase, Elk. The Elk gene, B1, encode a protein which is
XX CC a member of the Eph subfamily of protein tyrosine kinases. The Elk
XX CC product is very similar to two other receptor-like tyrosine kinases,
XX CC eph and eck. Lambda-BI-Elk may be used in the production of
XX CC phosphorylated exogenous protein along with a further vector encoding
XX CC the desired exogenous protein. These plasmid may be used to produce
XX CC phosphorylated proteins in host cells which have no intrinsic capacity
XX CC for phosphorylation, eg. bacteria. The system may be used for the
XX CC expression of the phosphorylated kinase insert domain of a growth
XX CC factor receptor kinase eg. platelet-derived growth factor receptor.
XX
XX SO Sequence 4281 BP; 1075 A; 1170 C; 1121 G; 915 T; 0 other;

Alignment Scores:
Pred. No.: 0.874 Length: 4281
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 14 Gaps: 0

US-09-397-967-16 (1-1099) x AA053471 (1-4281)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2587 GTGCACCGGGACCTGCTGTAGAACATCTGTGTC 2622

RESULT 97
AA092641
ID AA092641 standard; CDNA to mRNA: 4290 BP.
XX
XX AC AA092641;
XX
XX 06-MAR-1996 (first entry)
XX
XX Human non-differentiated blood cell tyrosine kinase receptor.
XX DE Human: tyrosine kinase; receptor; non-differentiated; blood cell;
XX KW detection; cancer; bone marrow; transplant; treatment; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT CDS 410..3373

```

```

FT      sig_peptide          /product= receptor-type tyrosine kinase protein
FT      FT                  /tag= C
FT      mat_peptide         450..3370
FT      FT                  /*tag= d
XX      JP08188596-A..
XX      PD        23-JUL-1996.
XX      PF        13-JAN-1995;    95JP-0003677.
XX      PE        09-NOV-1994;    94JP-0275411.
XX      PR        19-OCT-1994;    94JP-0253848.
XX      PA        (ASAH ) ASAH1 KASEI KOGYO KK.
XX      XA        WPI: 1996-388601/39.
XX      DR        P-PSDB; AAW06335.
XX      PT        New ligand for receptor type tyrosine kinase - has mol.wt. 22-25
XX      PI        kDa; dalton(s) and is positive for Coomassie and PAS staining
XX      PS        Example 16; Page 39-43; 51pp; Japanese.
CC      CC        This sequence represents the coding sequence for the full length
CC      CC        receptor-type tyrosine kinase protein. The encoded protein sequence is
CC      CC        recognised by a protein of the invention. The proteins of the invention
CC      CC        have a molecular weight of 23500 (plus or minus 1500) Da, and are
CC      CC        positive for Coomassie staining and PAS staining. The N-terminals of two
CC      CC        proteins of the invention (see AAW06333 and AAW06334) is represented by
CC      CC        AAW06332. The proteins of the invention bind to the fragment of the
CC      CC        annotated protein represented by AAW06330. The protein is a new ligand of
CC      CC        recombinant-type tyrosine kinases, and can be prepared by standard
CC      CC        recombinant techniques.
SQ      SQ        Sequence 4290 BP; 824 A; 1310 C; 1342 G; 809 T; 5 other:

Alignment Scores:
pred. No.:       0.876           Length:       4290
Score:           12.00           Matches:     12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match:     1.09%           Indels:      0
DB:              17             Gaps:        0

US-09-397-967-16 (1-1099) x AAT42593 (1-4290)
CY      CY      941 ValHisArgAspLeuAlaIalaIrgAsnIleuVal 952
                |||||
Db      Db      2618 GTCCACCGAGACTGGCTGCTGCACAACCTCAGTC 2653

RESULT 99
AAT18394 ID AAT18394 standard; cDNA to mRNA; 4290 BP.
XX      XX      AAT18394.
XX      AC      18-OCT-1996 (first entry)
XX      DT      Receptor type tyrosine kinase (TK) coding sequence.
DE      DE      KW Receptor type tyrosine kinase; TK; ligand; differentiation;
KW      KW      haematopoietic stem cell; tyrosine; bone marrow; leukaemia; ss.
XX      XX      OS Homo sapiens.
XX      XA      Key Location/Qualifiers
XX      FH      CDS 410..3373
XX      FT      /*tag= a
XX      FT      /product= Receptor type tyrosine kinase.
XX      FT      410..454
XX      FT      sig_peptide
XX      FT      /*tag= b

```

```

FT mat-peptide 455..3370
FT /*tag= c
PN WO9611212-A1.
XX 18-APR-1996.
PD
XX
XX 09-OCT-1995; 95WO-JP02069.
XX
XX 22-DEC-1994; 94JP-0320712.
PR 07-OCT-1994; 94JP-0244433.
PR 26-OCT-1994; 94JP-0262882.
XX
XX (ASAH ) ASAH KASEI KOGYO KK.
XX
XX Ohno M, Sakano S;
XX
XX WPI: 1996-209809/21.
DR P-PSDB; AAR94652, AAR94653.
XX
XX Lysand peptide binding to receptor-type tyrosine kinase - enhances
PT intracellular tyrosine phosphorylation, useful for investigation of
XX undifferentiated blood cell behaviour
XX
XX Claim 1; Page 154-161; 193pp; Japanese.
XX
XX A ligand polypeptide which binds to the extracellular part of a
CC specific receptor-type tyrosine kinase and induces phosphorylation
CC of tyrosine within the cell can be used in the study of the
CC differentiation of blood cells such as the haematopoietic stem
CC cells; of disease processes such as leukemia, and of the biology of
CC bone marrow transplantation. The ligand plays a role in the
CC of undifferentiation process and the specific ligand target is expressed
XX in undifferentiated blood cells.
XX
SQ Sequence 4290 BP; 824 A; 1323 C; 1329 G; 809 T; 5 other;
XX
Alignment Scores:
Pred. No.: 0.876 Length: 4290
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 17 Gaps: 0
US-09-397-967-16 (1-1099) x AAT18394 (1-4290)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
DB 2618 GTCCACCGAGACCTGCTGCTGCACATCTAGTC 2653
RESULT 100
AAT51235
ID AAT51235 standard; cDNA to mRNA; 4290 BP.
XX
AC AAT51235;
XX
XX 21-MAR-1997 (first entry)
XX
XX Receptor-type tyrosine kinase coding sequence.
XX
XX Receptor-type tyrosine kinase; LPM medium; ligand; autophosphorylation;
KW insect cell; animal cell; growth promoter; undifferentiated cell;
KW insulin; interleukin; fibroblast growth factor; hepatocyte growth factor;
KW nerve growth factor; interferon-gamma; tumour necrosis factor; inducer;
KW ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 410..3373
XX /*tag= a
XX /*product= receptor-type tyrosine kinase

```

```

FT sig-peptide 410..454
FT /*tag= b
FT mat-peptide 455..3370
FT /*tag= c
XX
XX JP08280385-A.
XX
XX 29-OCT-1996.
PD
XX
XX 13-APR-1995; 95JP-0087878.
PF
XX
XX 13-APR-1995; 95JP-0087878.
PR
XX
XX (ASAH ) ASAH KASEI KOGYO KK.
XX
XX WPI: 1997-014848/02.
DR P-PSDB; AAM1304.
XX
XX LPM cell culture medium - for culturing cells which produce
PT receptor-type tyrosine kinase ligand
XX
XX Claim 2; Page 34-38; 45pp; Japanese.
XX
XX This sequence represents the coding sequence for human receptor-type
CC tyrosine kinase. The protein encoded by this sequence is included in the
CC LPM medium of the invention. The medium is used to culture cells capable
CC of producing a ligand for receptor-type tyrosine kinase. The ligand
CC serves as a physiologically active substance to induce enzyme activity
CC and autophosphorylation of tyrosine kinase, which is involved in the
CC development and differentiation of animal and insect cells. The ligand
CC can also be used as in a pharmaceutical preparation for promoting growth
CC of undifferentiated cells. The medium may also include an inducer, such
CC as insulin, an interleukin, a fibroblast growth factor, or
XX interferon-gamma.
XX
SQ Sequence 4290 BP; 824 A; 1310 C; 1342 G; 809 T; 5 other;
XX
Alignment Scores:
Pred. No.: 0.876 Length: 4290
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 1
DB: 18 Gaps: 0
US-09-397-967-16 (1-1099) x AAT51235 (1-4290)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleuVal 952
DB 2618 GTCCACCGAGACCTGCTGCTGCACATCTAGTC 2653

```

Search completed: April 28, 2003, 19:08:25  
 Job time : 695 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: April 28, 2003, 21:08:40 ; Search time 267 Seconds

(without alignments)  
4478.852 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 1099

Sequence: 1 MAPSETPLIPORSCSLSS.....RPAFATLSPQDPLMRGRPG 1099

Scoring table:

OLIGO	Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0	
Egapop 6.0	Egapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 709820 seqs, 544064369 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1417911

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 150 summaries

Command line parameters:

```

-MODEL=frame+_p2n_model -DEV=xlp
-O=/cgn2_1/USPRO_spool/US09397967/runat_26042003_182929_11324/app_query.fasta_1.1287
-DB=Published Applications NA -OFMT=fastap -SUFFIX=oligo.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -STAR=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=150 -DOCALIGN=200 -THR.SCORE=quality -THR.MIN=1
-ALIGN=100 -MODE=LOCAL -OUTEXT=pio -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09397967@CON_1.1.175.@runat_26042003_182929_11324
-NCP=6 -ICPU=3 -NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT
-DSPLOCK=100 -LONGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60
-XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

```

Database: Published Applications\_NA.\*

```

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCIT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCITUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5.3	203	10	US-09-728-446-1413	Sequence 1413, Ap
2	4.6	439	9	US-09-918-995-11571	Sequence 11571, A
3	2.5	438	9	US-09-918-995-10879	Sequence 10879, A
4	1.6	364	10	US-09-771-161A-17	Sequence 17, Appl

5	1.1	151	10	US-09-982-610-11	Sequence 11, Appl
12	1.1	294	10	US-09-864-761-21086	Sequence 21086, A
6	1.1	367	10	US-09-864-761-4335	Sequence 4335, Ap
7	1.1	372	10	US-09-878-574-1581	Sequence 1581, Ap
8	1.1	477	9	US-09-918-995-2066	Sequence 2066, Ap
9	1.1	917	9	US-09-991-936-1577	Sequence 1577, Ap
10	1.1	2000	9	US-09-977-260-1	Sequence 1, Appl1
11	1.1	2000	9	US-09-977-261-1	Sequence 1, Appl1
12	1.1	3107	10	US-09-977-269-1	Sequence 1, Appl1
13	1.1	3348	10	US-09-964-456-1128	Sequence 1128, Ap
14	1.1	3370	10	US-09-982-610-34	Sequence 34, Appl
15	1.1	3393	10	US-09-967-768A-144	Sequence 144, Appl
16	1.1	3583	9	US-09-922-138-3	Sequence 3, Appl
17	1.1	3583	9	US-10-245-103-103	Sequence 103, Appl
18	1.1	3583	9	US-10-245-107-103	Sequence 103, Appl
19	1.1	3583	9	US-10-245-143-103	Sequence 103, Appl
20	1.1	3583	9	US-10-245-171-103	Sequence 103, Appl
21	1.1	3583	9	US-10-245-851-103	Sequence 103, Appl
22	1.1	3583	9	US-10-245-883-103	Sequence 103, Appl
23	1.1	3583	9	US-10-237-535-103	Sequence 103, Appl
24	1.1	3583	9	US-10-238-183-103	Sequence 103, Appl
25	1.1	3583	9	US-10-238-283-103	Sequence 103, Appl
26	1.1	3583	9	US-10-238-370-103	Sequence 103, Appl
27	1.1	3583	9	US-10-245-055-103	Sequence 103, Appl
28	1.1	3583	9	US-10-245-147-103	Sequence 103, Appl
29	1.1	3583	9	US-10-245-730-103	Sequence 103, Appl
30	1.1	3583	9	US-10-245-739-103	Sequence 103, Appl
31	1.1	3583	9	US-10-246-210-103	Sequence 103, Appl
32	1.1	3633	10	US-09-922-138-1	Sequence 1, Appl1
33	1.1	3643	10	US-09-921-771-4	Sequence 4, Appl1
34	1.1	3969	10	US-09-982-610-23	Sequence 23, Appl
35	1.1	4698	10	US-09-895-652-1	Sequence 1, Appl1
36	1.1	5816	10	US-09-205-658-11	Sequence 11, Appl
37	1.1	5816	10	US-09-844-353A-11	Sequence 11, Appl
38	1.1	147	10	US-09-982-610-7	Sequence 7, Appl1
39	1.0	3018	8	US-09-938-812A-669	Sequence 669, Appl
40	1.0	4165	8	US-08-578-684-1	Sequence 1, Appl1
41	1.0	510	9	US-09-796-632-8268	Sequence 8268, Ap
42	0.9	516	9	US-10-101-464A-355	Sequence 355, Appl
43	0.9	539	9	US-09-796-692-8255	Sequence 8255, Ap
44	0.9	556	9	US-09-796-692-8283	Sequence 7883, Ap
45	0.9	556	9	US-09-796-692-8285	Sequence 8025, Ap
46	0.9	564	9	US-09-796-692-9025	Sequence 7025, Ap
47	0.9	589	9	US-09-796-692-8630	Sequence 8630, Ap
48	0.9	597	9	US-09-796-692-7976	Sequence 7976, Ap
49	0.9	597	9	US-09-796-692-8273	Sequence 8273, Ap
50	0.9	601	9	US-09-796-692-8047	Sequence 8047, Ap
51	0.9	605	9	US-09-796-692-8927	Sequence 8927, Ap
52	0.9	608	9	US-09-796-692-8132	Sequence 8132, Ap
53	0.9	610	9	US-09-796-692-8085	Sequence 8085, Ap
54	0.9	645	9	US-09-796-692-7538	Sequence 7538, Ap
55	0.9	869	9	US-10-101-464A-368	Sequence 368, Appl
56	0.9	1021	9	US-10-101-464A-218	Sequence 218, Appl
57	0.9	1128	9	US-10-101-464A-343	Sequence 343, Appl
58	0.9	2040	9	US-10-101-464A-872	Sequence 872, Appl
59	0.9	2052	9	US-09-938-842A-1497	Sequence 1497, Ap
60	0.9	2136	9	US-09-938-842A-1480	Sequence 1480, Ap
61	0.9	2820	10	US-09-924-859A-4	Sequence 4, Appl1
62	0.9	3120	10	US-09-982-610-22	Sequence 22, Appl
63	0.9	3270	10	US-09-866-510-1	Sequence 1, Appl1
64	0.9	3270	10	US-09-866-510-3	Sequence 3, Appl1
65	0.9	3270	10	US-09-866-510-9	Sequence 9, Appl1
66	0.9	3270	10	US-09-866-510-13	Sequence 13, Appl
67	0.9	3321	10	US-09-866-510-15	Sequence 15, Appl
68	0.9	3321	10	US-09-866-510-17	Sequence 17, Appl
69	0.9	3321	10	US-09-866-510-21	Sequence 21, Appl
70	0.9	3321	10	US-09-866-510-21	Sequence 21, Appl
71	0.9	3453	10	US-09-872-136-1	Sequence 1, Appl1
72	0.9	3453	10	US-09-872-136-3	Sequence 3, Appl1
73	0.9	3501	10	US-09-919-408-3	Sequence 3, Appl1
74	0.9	4054	10	US-09-872-136-3	Sequence 3, Appl1
75	0.9	4054	10	US-09-955-363-35	Sequence 35, Appl1
76	0.9	4071	9	US-10-022-939-1	Sequence 1, Appl1
77	0.9				

78	10	0.9	4071	9	US-10-100-045A-1	Sequence 1, Appl1
79	10	0.9	4111	9	US-09-375-208-1	Sequence 1, Appl1
80	10	0.9	4425	10	US-09-982-610-31	Sequence 31, Appl1
81	10	0.9	4465	9	US-09-955-363-1	Sequence 1, Appl1
82	10	0.9	4989	9	US-09-870-759-119	Sequence 119, Appl1
83	10	0.9	5084	10	US-09-954-531-1383	Sequence 1383, Ap
84	10	0.9	5084	10	US-09-967-768A-277	Sequence 277, App
85	10	0.9	5406	10	US-09-919-408-5	Sequence 5, Appl1
86	10	0.9	5406	10	US-09-872-136-5	Sequence 5, Appl1
87	10	0.9	5470	10	US-09-766-678-1	Sequence 1, Appl1
88	10	0.9	6378	10	US-09-919-497-40	Sequence 40, Appl1
89	10	0.9	6412	10	US-09-769-987-1	Sequence 1, Appl1
90	10	0.9	6827	10	US-09-982-610-17	Sequence 17, Appl1
91	10	0.9	9108	10	US-09-982-610-45	Sequence 45, Appl1
92	9	0.8	3352	10	US-09-960-352-10875	Sequence 10875, A
93	9	0.8	355	10	US-09-960-352-1177	Sequence 1177, Ap
94	9	0.8	363	9	US-10-101-464A-332	Sequence 332, App
95	9	0.8	366	10	US-09-960-352-10878	Sequence 10878, A
96	9	0.8	391	10	US-09-960-352-2603	Sequence 2603, Ap
97	9	0.8	392	10	US-09-960-352-4293	Sequence 4293, Ap
98	10	0.8	401	10	US-09-960-352-13765	Sequence 13765, A
99	9	0.8	402	10	US-09-960-352-14442	Sequence 14442, A
100	9	0.8	415	9	US-09-918-995-16302	Sequence 16302, A
101	9	0.8	415	10	US-09-960-352-10423	Sequence 10423, A
102	9	0.8	422	10	US-09-960-352-6936	Sequence 6936, Ap
103	9	0.8	434	10	US-09-960-352-5227	Sequence 5227, Ap
104	9	0.8	462	10	US-09-864-761-20164	Sequence 20164, A
105	9	0.8	463	9	US-10-102-806-228	Sequence 228, App
106	9	0.8	463	9	US-09-918-995-23959	Sequence 23959, A
107	9	0.8	480	9	US-10-060-036-4148	Sequence 4148, Ap
108	9	0.8	482	9	US-09-796-662-6991	Sequence 6991, Ap
109	9	0.8	486	9	US-10-101-464A-378	Sequence 378, App
110	9	0.8	513	9	US-09-918-995-14461	Sequence 14461, A
111	9	0.8	753	10	US-09-910-943-636	Sequence 636, App
112	9	0.8	903	10	US-09-815-242-4130	Sequence 4130, Ap
113	9	0.8	1620	9	US-10-118-984-3	Sequence 3, Appl1
114	9	0.8	1620	10	US-09-728-721-3	Sequence 3, Appl1
115	9	0.8	1620	12	US-10-105-931-3	Sequence 3, Appl1
116	9	0.8	1755	9	US-09-930-125-6	Sequence 6, Appl1
117	9	0.8	1767	9	US-09-930-125-4	Sequence 4, Appl1
118	9	0.8	1773	9	US-09-930-125-7	Sequence 7, Appl1
119	9	0.8	1806	9	US-09-930-125-5	Sequence 5, Appl1
120	9	0.8	1931	9	US-10-118-984-1	Sequence 1, Appl1
121	9	0.8	1931	10	US-09-748-537-2	Sequence 2, Appl1
122	9	0.8	1931	10	US-09-728-721-1	Sequence 1, Appl1
123	9	0.8	1931	12	US-10-133-780-2	Sequence 2, Appl1
124	9	0.8	1931	12	US-10-105-931-1	Sequence 1, Appl1
125	9	0.8	2184	10	US-09-728-952-82	Sequence 82, Appl1
126	9	0.8	2187	9	US-09-954-531-188	Sequence 188, App
127	9	0.8	2338	10	US-09-925-302-90	Sequence 90, Appl1
128	9	0.8	2437	9	US-09-158-722-3	Sequence 3, Appl1
129	9	0.8	2454	10	US-09-964-824A-120	Sequence 120, App
130	9	0.8	2627	10	US-09-964-824-313	Sequence 313, App
131	9	0.8	2709	10	US-09-925-301-173	Sequence 173, App
132	9	0.8	3454	12	US-10-044-090-48	Sequence 48, Appl1
133	9	0.8				

```

ALIGNMENTS

RESULT 1
US-09-728-446-1413
; Sequence 1413, Application US/09728446
; Patent No. US20020081668A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; TITLE OF INVENTION: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020081668A1el Murine Polynucleotide Sequences
; FILE REFERENCE: LEX-0101-USA
; CURRENT APPLICATION NUMBER: US/09/728, 446
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,270
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 1461
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1413
; LENGTH: 203
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-728-446-1413

Alignment Scores:
Pred. No.:          3.47e-50      Length:        203
Score:              58.00         Matches:        58
Percent Similarity: 100.00%       Conservative:   0
Best local Similarity: 100.00%    Mismatches:    0
Query Match:        5.28%         Indels:         0
DB:                 10            Gaps:           0

US-09-397-967-16 (1-1099) x US-09-728-446-1413 (1-203)
Qy      1  MetaLAPProSeSerGIuGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
      |  |||||
Db      18  AMGGACCTCCAGTGAGGAGAGACACTGTGATCCCTCAAGCGCTTGCAGCCTCATCC 77
Qy      21  SerGuAlaGlyAlaLeuHisValLeuLeuP-ProProlArgGlyProGlyProProGlnArg 40
      |  |||||
Db      78  TCAGAGGAGGAGGCCCTGCATGTGCTCTTCCTCCCGGGGAGCTCGGGCTCCCGCAGCGA 137
Qy      41  LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValaArgAlaala 58
      |  |||||
Db     138  TTGTATTCTCTTTTGGGGACTACTTGGCTAGAGATTATGTGTGGAGACTGCGC 191

RESULT 2
US-09-918-995-11571
; Sequence 11571, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: FROM VARIOUS CGNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11571
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(439)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-11571

```



```
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 151 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-982-610-11

Alignment Scores:
Pred. No.: 0.0027 Length: 151
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-982-610-11 (1-151)
Qy 941 ValHisArgSplLeuAlaAlaArgAsnIleLeuVal 952
Db 1 GTGCACAGGATCTCGCGGCTCGGACATCTCTCTC 36

RESULT 6
US-09-864-761-21086
; Sequence 21086, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomlca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
```

```
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 21086
LENGTH: 294
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC010899.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.8
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.4
OTHER INFORMATION: SWISSPROT HIT: Q91694, EVALUE 2.00e-35
OTHER INFORMATION: NT HIT: g14758279, EVALUE 1.00e-115
OTHER INFORMATION: EST_HUMAN HIT: AW965297.1, EVALUE 2.00e-26
US-09-864-761-21086

Alignment Scores:
Pred. No.: 0.0049 Length: 294
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-864-761-21086 (1-294)
Qy 941 ValHisArgSplLeuAlaAlaArgAsnIleLeuVal 952
Db 91 GTGCATCGTATCTGCGCGACGAGACATCTCTG 126

RESULT 7
US-09-864-761-4335
; Sequence 4335, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomlca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
```



```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4335
; LENGTH: 367
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010899.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HELI, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.4
; US-09-864-761-4335

Alignment Scores:
Pred. No.: 0.00597 Length: 367
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-864-761-4335 (1-367)

QY 941 VALHISARGSPLEUALAALARGASNIIEUVAL 952
DB 289 GTGCATCGTGCATCTGCGCCGACGAAACATCTGCTG 324

RESULT 8
; Sequence 1581, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 1581
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: L1B3028-036-01-B1-B10
; US-09-878-574-1581

Alignment Scores:
```

```
Pred. No.: 0.00605 Length: 372
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-878-574-1581 (1-372)

QY 1003 SERSAPVALTRPSRPHGLYVALLEUITYRGU 1014
DB 329 AGTGATGTGTGGAGTTTGGGGTGTCTATATGAA 364

RESULT 9
; Sequence 2066, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2066
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)-(477)
; OTHER INFORMATION: n = A,T,C or G
; US-09-918-995-2066

Alignment Scores:
Pred. No.: 0.00755 Length: 477
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-918-995-2066 (1-477)

QY 941 VALHISARGSPLEUALAALARGASNIIEUVAL 952
DB 200 GTGCACCGCGACCTGCTGCTGCAACATCTTGTG 235

RESULT 10
; Sequence 1577, Application US/09991936
; Publication No. US20030073627A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Miszewski, Nancy
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1577
```

```
; LENGTH: 917
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-1577

Alignment Scores:
Pred. No.: 0.0136      Length: 917
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%              Indels: 0
DB: 9                      Gaps: 0

US-09-397-967-16 (1-1099) x US-09-991-936-1577 (1-917)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
|||||
Db 169 GTCCATCGAGATTGGCTGCCAGGAATATCTGTG 204

RESULT 11
US-09-977-260-1
; Sequence 1, Application US/09977260
; Publication No. US20020192790A1
; GENERAL INFORMATION:
; APPLICANT: ULRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,260
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (258)..(1778)
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 1
US-09-977-260-1

Alignment Scores:
Pred. No.: 0.0272      Length: 2000
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%              Indels: 0
DB: 9                      Gaps: 0

US-09-397-967-16 (1-1099) x US-09-977-260-1 (1-2000)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
|||||
Db 1302 GTGCACCGCGACCTGGCGCCGCCACACATCTGTG 1337

RESULT 12
US-09-977-261-1
; Sequence 1, Application US/09977261
; Publication No. US20030054527A1
; GENERAL INFORMATION:
; APPLICANT: ULRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1259
; CURRENT APPLICATION NUMBER: US/09/977,261
; CURRENT FILING DATE: 2001-10-16
```

```
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (258)..(1778)
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 1
US-09-977-261-1

Alignment Scores:
Pred. No.: 0.0272      Length: 2000
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%              Indels: 0
DB: 10                     Gaps: 0

US-09-397-967-16 (1-1099) x US-09-977-261-1 (1-2000)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
|||||
Db 1302 GTGCACCGCGACCTGGCGCCGCCACACATCTGTG 1337

RESULT 13
US-09-977-269-1
; Sequence 1, Application US/09977269
; Patent No. US20020082037A1
; GENERAL INFORMATION:
; APPLICANT: ULRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,269
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (258)..(1778)
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 1
US-09-977-269-1

Alignment Scores:
Pred. No.: 0.0272      Length: 2000
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%              Indels: 0
DB: 10                     Gaps: 0

US-09-397-967-16 (1-1099) x US-09-977-269-1 (1-2000)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
|||||
Db 1302 GTGCACCGCGACCTGGCGCCGCCACACATCTGTG 1337

RESULT 14
```

US-09-954-456-1128  
; Sequence 1128, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1128  
; LENGTH: 3107  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-1128

Alignment Scores:  
Pred. No.: 0.0404 Length: 3107  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservatave: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-954-456-1128 (1-3107)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
DB 2260 GTGCATCGTGATCTGCGCCGACGAGACATCTGCTG 2295

RESULT 15  
US-09-982-610-34  
; Sequence 34, Application US/09982610  
; Patent No. US20020146420A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; Bennett, Brian D.  
; Goeddel, David  
; Lee, James M.  
; Matthews, William  
; Tsai, Siao Ping  
; Wood, William I.  
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 Inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/982,610  
FILING DATE: 17-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/446,648  
FILING DATE: 1996-MAY-23  
APPLICATION NUMBER: 08/222616  
FILING DATE: 04-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0821P3PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3348 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 34:  
US-09-982-610-34

Alignment Scores:  
Pred. No.: 0.0432 Length: 3348  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservatave: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-982-610-34 (1-3348)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
DB 2227 GTGCATCGTGATCTGCGCCGACGAGACATCTGCTG 2262

RESULT 16  
US-09-967-768A-144  
; Sequence 144, Application US/09967768A  
; Patent No. US20020150877A1  
; GENERAL INFORMATION:  
; APPLICANT: Augustus, Meena  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign  
; FILE REFERENCE: 689290-72  
; CURRENT APPLICATION NUMBER: US/09/967,768A  
; FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,109  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,034  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,111  
; PRIOR FILING DATE: 2000-09-28  
; NUMBER OF SEQ ID NOS: 325  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 144  
; LENGTH: 3370  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-967-768A-144

Alignment Scores:  
Pred. No.: 0.0435 Length: 3370  
Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0  
DB: 10

US-09-397-967-16 (1-1099) x US-09-967-768A-144 (1-3370)

QY 941 VALHISARGSPLEUALAALARGASNILEUVAL 952  
DB 2353 GTCCACGGGAGCTGCTGCCAGAAACATCTTGTTG 2388

RESULT 17  
US-09-922-138-3  
; Sequence 3, Application US/09922138  
; Patent No. US20020061574A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel  
; APPLICANT: Silos-Santiago, Immaculada  
; TITLE OF INVENTION: 1658, 14223, AND 16002, NOVEL HUMAN  
; TITLE OF INVENTION: KINASES AND USES THEREFOR  
; FILE REFERENCE: 38155-20030.00  
; CURRENT APPLICATION NUMBER: US/09/922.138  
; CURRENT FILING DATE: 2001-11-28  
; PRIOR APPLICATION NUMBER: US 60/229,299  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 3393  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-922-138-3

Alignment Scores:  
Pred. No.: 0.0437 Length: 3393  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0  
DB: 10

US-09-397-967-16 (1-1099) x US-09-922-138-3 (1-3393)

QY 941 VALHISARGSPLEUALAALARGASNILEUVAL 952  
DB 2665 GTTCATCGAGACCTAGCGGCTCGAATATACTGTC 2700

RESULT 18  
US-10-245-103-103  
; Sequence 103, Application US/10245103  
; Publication No. US20030068778A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3630R1C112  
; CURRENT APPLICATION NUMBER: US/10/245.103  
; CURRENT FILING DATE: 2002-09-17  
; PRIOR APPLICATION NUMBER: 10/197942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114

; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/065027  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/086478  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090689  
; PRIOR FILING DATE: 1998-06-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 116  
; SEQ ID NO 103  
; LENGTH: 3583  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-245-103-103

Alignment Scores:  
Pred. No.: 0.0459 Length: 3583  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0  
DB: 9

US-09-397-967-16 (1-1099) x US-10-245-103-103 (1-3583)

QY 941 VALHISARGSPLEUALAALARGASNILEUVAL 952  
DB 2633 GTTCATCGAGACCTAGCGGCTCGAATATACTGTC 2668

RESULT 19  
US-10-245-107-103  
; Sequence 103, Application US/10245107  
; Publication No. US20030068779A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Watanabe, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3630R1C71  
; CURRENT APPLICATION NUMBER: US/10/245.107  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: 10/197942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/065027  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/086478  
; PRIOR FILING DATE: 1998-05-22

```

; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SEQ ID NO 103
; LENGTH: 3583
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-107-103

Alignment Scores:
Pred. No.: 0.0459      Length: 3583
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%              Indels: 0
DB: 9                    Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-107-103 (1-3583)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2633 GTTCATCGAGACCTAGCGGCTCGGAATATATCTGCTC 2668

RESULT 20
US-10-245-143-103
; Sequence 103, Application US/10245143
; Publication No. US20030068780A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watande, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C90
; CURRENT APPLICATION NUMBER: US/10/245,143
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```

; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 103
; LENGTH: 3583
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-143-103

Alignment Scores:
Pred. No.: 0.0459      Length: 3583
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%              Indels: 0
DB: 9                    Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-143-103 (1-3583)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2633 GTTCATCGAGACCTAGCGGCTCGGAATATATCTGCTC 2668

RESULT 21
US-10-245-771-103
; Sequence 103, Application US/10245771
; Publication No. US20030068781A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watande, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C98
; CURRENT APPLICATION NUMBER: US/10/245,771
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SEQ ID NO 103
; LENGTH: 3583
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-771-103

Alignment Scores:
Pred. No.: 0.0459      Length: 3583
```

Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 9 Caps: 0

US-09-397-967-16 (1-1099) x US-10-245-771-103 (1-3583)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952

DB 2633 GTTCATCGAGACCTAGCGGCTCGGAATATACTGCTC 2668

RESULT 22

US-10-245-851-103  
; Sequence 103, Application US/10245851  
; Publication No. US20030068782A1

; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Matande, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3630R1C93  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: US/10/245, 851  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 10/197942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/065027  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/086478  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090689  
; PRIOR FILING DATE: 1998-06-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 116  
; SEQ ID NO 103  
; LENGTH: 3583  
; TYPE: DNA  
; ORGANISM: Homo Sapien

US-10-245-851-103

Alignment Scores:

Pred. No.: 0.0459 Length: 3583  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-851-103 (1-3583)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952

DB 2633 GTTCATCGAGACCTAGCGGCTCGGAATATACTGCTC 2668

RESULT 23

US-10-245-883-103  
; Sequence 103, Application US/10245883  
; Publication No. US20030068783A1

; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin  
; APPLICANT: Eaton, Dan  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Phillippe  
; APPLICANT: Matande, Colin  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; APPLICANT: Fong, Sherman  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3630R1C70  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: US/10/245, 883  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 10/197942  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/059114  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/063046  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: 60/065027  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/086478  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090689  
; PRIOR FILING DATE: 1998-06-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 116  
; SEQ ID NO 103  
; LENGTH: 3583  
; TYPE: DNA  
; ORGANISM: Homo Sapien

US-10-245-883-103

Alignment Scores:

Pred. No.: 0.0459 Length: 3583  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-883-103 (1-3583)

OY 941 VALHISARGASPLEUALAALARGASNILEUVAL 952

DB 2633 GTTCATCGAGACCTAGCGGCTCGGAATATACTGCTC 2668

RESULT 24

US-10-237-535-103  
; Sequence 103, Application US/10237535  
; Publication No. US20030073188A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin

APPLICANT: Eaton, Dan  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Goddard, Audrey  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin  
APPLICANT: Smith, Victoria  
APPLICANT: Stephan, Jean-Phillippe  
APPLICANT: Watanabe, Colin  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
APPLICANT: Fong, Sherman  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P363OR1C3  
CURRENT APPLICATION NUMBER: US/10/237,535  
CURRENT FILING DATE: 2002-09-06  
PRIOR APPLICATION NUMBER: 10/17942  
PRIOR FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: 60/059114  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/063046  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/065027  
PRIOR FILING DATE: 1997-11-10  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/086478  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090689  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/091358  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/099803  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: 60/106932  
PRIOR FILING DATE: 1998-11-03  
PRIOR APPLICATION NUMBER: 60/115554  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/119342  
PRIOR FILING DATE: 1999-02-09  
PRIOR APPLICATION NUMBER: 60/123957  
PRIOR FILING DATE: 1999-03-12  
PRIOR APPLICATION NUMBER: 60/123972  
PRIOR FILING DATE: 1999-03-11  
PRIOR APPLICATION NUMBER: 60/127372  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/131271  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/133459  
PRIOR FILING DATE: 1999-05-11  
PRIOR APPLICATION NUMBER: 60/135725  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/135729  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/135750  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/138385  
PRIOR FILING DATE: 1999-06-09  
PRIOR APPLICATION NUMBER: 60/140653  
PRIOR FILING DATE: 1999-06-22  
PRIOR APPLICATION NUMBER: 60/141037  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: 60/144732  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/144758

PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/144790  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/145228  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/145698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: 60/146222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: 60/146843  
PRIOR FILING DATE: 1999-08-03  
PRIOR APPLICATION NUMBER: 60/148188  
PRIOR FILING DATE: 1999-08-10  
PRIOR APPLICATION NUMBER: 60/148513  
PRIOR FILING DATE: 1999-08-12  
PRIOR APPLICATION NUMBER: 60/149327  
PRIOR FILING DATE: 1999-08-17  
PRIOR APPLICATION NUMBER: 60/149395  
PRIOR FILING DATE: 1999-08-17  
PRIOR APPLICATION NUMBER: 60/150114  
PRIOR FILING DATE: 1999-08-20  
PRIOR APPLICATION NUMBER: 60/151700  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: 60/151734  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: 60/162506  
PRIOR FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: 60/170262  
PRIOR FILING DATE: 1999-12-09  
PRIOR APPLICATION NUMBER: 60/177118  
PRIOR FILING DATE: 2000-01-20  
PRIOR APPLICATION NUMBER: 60/179851  
PRIOR FILING DATE: 2000-02-02  
PRIOR APPLICATION NUMBER: 60/180921  
PRIOR FILING DATE: 2000-02-08  
PRIOR APPLICATION NUMBER: 60/187202  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: 60/198587  
PRIOR FILING DATE: 2000-04-18  
PRIOR APPLICATION NUMBER: 60/199614  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 60/206330  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/206368  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/209832  
PRIOR FILING DATE: 2000-06-05  
PRIOR APPLICATION NUMBER: 60/218371  
PRIOR FILING DATE: 2000-07-13  
PRIOR APPLICATION NUMBER: 60/222695  
PRIOR FILING DATE: 2000-08-02  
PRIOR APPLICATION NUMBER: 60/229896  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/230621  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: 60/232887  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: 60/235147  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 60/261878  
PRIOR FILING DATE: 2001-01-12  
PRIOR APPLICATION NUMBER: 60/261910  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: 60/261939  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: 60/262150  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: 60/264395  
PRIOR FILING DATE: 2001-01-25  
PRIOR APPLICATION NUMBER: 60/266421  
PRIOR FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: 60/267623  
PRIOR FILING DATE: 2001-02-09

```
; PRIOR APPLICATION NUMBER: 60/274399
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/280982
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/282129
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/282199
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/290589
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: 09/180997
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: 09/267213
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423741
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: 09/709238
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/802706
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 09/872035
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/924419
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/927796
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/929404
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 09/931836
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 09/941992
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/946374
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 10/001054
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/081056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 10/119480
; PRIOR FILING DATE: 2002-04-09
```

## Alignment Scores:

Pred. No.:	0.0459	Length:	3583
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
	9	Gaps:	0

US-09-397-967-16 (1-1099) x US-10-237-535-103 (1-3583)

OY 941 ValHisArgAspLeuAlaIleArgAsnIleLeuVal 952

Db 2633 GTTCATCGAGACCTAGCGGCTCGGAAATATACTGCTC 2668

## RESULT 25

```
; Sequence 103, Application US/10238183
; Publication No. US20030073189A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
```

```
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watande, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C11
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091358
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/106932
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/115554
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119342
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/123957
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123972
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 60/127372
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/131271
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/134459
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/135725
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/135729
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/135750
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/138385
; PRIOR FILING DATE: 1999-06-09
; PRIOR APPLICATION NUMBER: 60/140653
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144732
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/144790
```



```

PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/282199
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/290589
PRIOR FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 09/267213
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/927796
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 09/931836
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 09/941992
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 10/001054
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 10/081056
PRIOR FILING DATE: 2002-02-20
PRIOR APPLICATION NUMBER: 10/119480
PRIOR FILING DATE: 2002-04-09

Alignment Scores:
Pred. No.: 0.0459
Score: 12.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.09%

DB: 9
Gaps: 0

US-09-397-967-16 (1-1099) x US-10-238-183-103 (1-3583)

QY 941 ValHisArgAspDeuAlaIlaIArgsnlleuVal 952
|||||
Db 2633 GTTCATCGAGACTACGCGCTCGAATATACGTC 2668

RESULT 26
; Sequence 103, Application US/10238283
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Elvatoroff, Ellen
; APPLICANT: Elvatoroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher

```

```

: APPLICANT: Gurney,Austin
: APPLICANT: Smith,Victoria
: APPLICANT: Stephan,Dean-Phillippe
: APPLICANT: Watanabe,Colin
: APPLICANT: Wood,William
: APPLICANT: Zhang,Zemin
: APPLICANT: Fong,Sherman
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3630R1C10
: CURRENT APPLICATION NUMBER: US/10/238,283
: CURRENT FILING DATE: 2002-09-09
: PRIOR APPLICATION NUMBER: 10/197942
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/059114
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/063046
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/065027
: PRIOR FILING DATE: 1997-11-10
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/086478
: PRIOR FILING DATE: 1998-05-22
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/090557
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090689
: PRIOR FILING DATE: 1998-06-25
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 116
: SEQ ID NO 103
: LENGTH: 3583
: TYPE: DNA
: ORGANISM: Homo Sapien
: US-10-238-283-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 9

US-09-397-967-16 (1-1099) x US-10-238-283-103 (1-3583)
QY 941 VALHISARGASPLEUAIAlaIaRgSnIleuVal 952
Db 2633 GTTCATCGAGACTTAGCGGCTCGGAAATATACTGCTC 2668

RESULT 27
US-10-238-370-103
: Sequence 103, Application US/10238370
: Publication No. US20030073191A1
: GENERAL INFORMATION:
: APPLICANT: Baker,Kevin
: APPLICANT: Eaton,Dan
: APPLICANT: Filvaroff,Ellen
: APPLICANT: Goddard,Audrey
: APPLICANT: Grimaldi,J. Christopher
: APPLICANT: Gurney,Austin
: APPLICANT: Smith,Victoria
: APPLICANT: Stephan,Dean-Phillippe
: APPLICANT: Watanabe,Colin
: APPLICANT: Wood,William
: APPLICANT: Zhang,Zemin
: APPLICANT: Fong,Sherman
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
```

```

: FILE REFERENCE: P3630R1C10
: CURRENT APPLICATION NUMBER: US/10/238,370
: CURRENT FILING DATE: 2002-09-09
: PRIOR APPLICATION NUMBER: 10/197942
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/059114
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/063046
: PRIOR FILING DATE: 1997-10-24
: PRIOR APPLICATION NUMBER: 60/065027
: PRIOR FILING DATE: 1997-11-10
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/086478
: PRIOR FILING DATE: 1998-05-22
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/090557
: PRIOR FILING DATE: 1998-06-24
: PRIOR APPLICATION NUMBER: 60/090689
: PRIOR FILING DATE: 1998-06-25
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 116
: SEQ ID NO 103
: LENGTH: 3583
: TYPE: DNA
: ORGANISM: Homo Sapien
: US-10-238-370-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 9

US-09-397-967-16 (1-1099) x US-10-238-370-103 (1-3583)
QY 941 VALHISARGASPLEUAIAlaIaRgSnIleuVal 952
Db 2633 GTTCATCGAGACTTAGCGGCTCGGAAATATACTGCTC 2668

RESULT 28
US-10-245-055-103
: Sequence 103, Application US/10245055
: Publication No. US20030073192A1
: GENERAL INFORMATION:
: APPLICANT: Baker,Kevin
: APPLICANT: Eaton,Dan
: APPLICANT: Filvaroff,Ellen
: APPLICANT: Goddard,Audrey
: APPLICANT: Grimaldi,J. Christopher
: APPLICANT: Gurney,Austin
: APPLICANT: Smith,Victoria
: APPLICANT: Stephan,Dean-Phillippe
: APPLICANT: Watanabe,Colin
: APPLICANT: Wood,William
: APPLICANT: Zhang,Zemin
: APPLICANT: Fong,Sherman
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3630R1C88
: CURRENT APPLICATION NUMBER: US/10/245,055
: CURRENT FILING DATE: 2002-09-16
: PRIOR APPLICATION NUMBER: 10/197942
: PRIOR FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/059114
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/063046
: PRIOR FILING DATE: 1997-10-24
```

```

; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 103
; LENGTH: 3583
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-055-103

Alignment Scores:
Pred. No.: 0.0459      Length: 3583
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%      Indels: 0
DB: 9                  Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-055-103 (1-3583)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2633 GTTCATCGAGACCTACGCGCTCGAATATCTGTC 2668

RESULT 29
US-10-245-147-103
; Sequence 103, Application US/10245147
; Publication No. US20030073193A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Collin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C72
; CURRENT APPLICATION NUMBER: US/10/245,147
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 103
; LENGTH: 3583
```

```

; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 103
; LENGTH: 3583
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-147-103

Alignment Scores:
Pred. No.: 0.0459      Length: 3583
Score: 12.00           Matches: 12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.09%      Indels: 0
DB: 9                  Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-147-103 (1-3583)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2633 GTTCATCGAGACCTACGCGCTCGAATATCTGTC 2668

RESULT 30
US-10-245-730-103
; Sequence 103, Application US/10245730
; Publication No. US20030073194A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Collin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C85
; CURRENT APPLICATION NUMBER: US/10/245,730
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 103
; LENGTH: 3583
```

```
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-730-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-730-103 (1-3583)

QY 941 VALHISARGASPLEUAAALARGASNIIEUVAL 952
DB 2633 GTTCATCGAGACCTAGCGGCTCGGAATATACGTGTC 2668

RESULT 31
US-10-245-739-103
; Sequence 103, Application US/10245739
; Publication No. US20030073195A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C86
; CURRENT APPLICATION NUMBER: US/10/245,739
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 103
; LENGTH: 3583
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-739-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0
```

```
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-245-739-103 (1-3583)

QY 941 VALHISARGASPLEUAAALARGASNIIEUVAL 952
DB 2633 GTTCATCGAGACCTAGCGGCTCGGAATATACGTGTC 2668

RESULT 32
US-10-246-210-103
; Sequence 103, Application US/10246210
; Publication No. US20030073196A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C121
; CURRENT APPLICATION NUMBER: US/10/246,210
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 103
; LENGTH: 3583
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-246-210-103

Alignment Scores:
Pred. No.: 0.0459 Length: 3583
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-246-210-103 (1-3583)

QY 941 VALHISARGASPLEUAAALARGASNIIEUVAL 952
DB 2633 GTTCATCGAGACCTAGCGGCTCGGAATATACGTGTC 2668
```



;; TITLE OF INVENTION: Treating Colon Cancer  
;; FILE REFERENCE: DEX-0211  
;; CURRENT APPLICATION NUMBER: US/09/895,652  
;; CURRENT FILING DATE: 2001-06-28  
;; PRIOR APPLICATION NUMBER: 60/214,515  
;; PRIOR FILING DATE: 2000-06-28  
;; NUMBER OF SEQ ID NOS: 39  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 1  
;; LENGTH: 4698  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: unsure  
;; LOCATION: (3921)  
US-09-895-652-1

Alignment Scores:  
Pred. No.: 0.0585 Length: 4698  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0

DB:  
US-09-397-967-16 (1-1099) x US-09-895-652-1 (1-4698)

OY 941 ValHsAtgApLeuAlAlArGsAnIleVal 952  
|||||  
Db 2237 GTTCACCGTGCCTGGCTGCCGCAACATCTCTGTC 2272

RESULT 37  
US-09-205-658-11

;; Sequence 11, Application US/09205658  
;; Patent No. US20010029617A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ruvkun, Gary  
;; APPLICANT: Ogg, Scott  
;; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR  
;; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS  
;; FILE REFERENCE: 00786/351004  
;; CURRENT APPLICATION NUMBER: US/09/205,658  
;; CURRENT FILING DATE: 1998-12-03  
;; EARLIER APPLICATION NUMBER: 08/857,076  
;; EARLIER FILING DATE: 1997-05-15  
;; EARLIER APPLICATION NUMBER: 08/888,534  
;; EARLIER FILING DATE: 1997-07-07  
;; EARLIER APPLICATION NUMBER: US98/10080  
;; EARLIER FILING DATE: 1998-05-15  
;; NUMBER OF SEQ ID NOS: 328  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 11  
;; LENGTH: 5816  
;; TYPE: DNA  
;; ORGANISM: Caenorhabditis elegans  
;; FEATURE:  
;; NAME/KEY: misc.feature  
;; LOCATION: (1)...(5816)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-205-658-11

Alignment Scores:  
Pred. No.: 0.0708 Length: 5816  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x US-09-205-658-11 (1-5816)

OY 1303 SerAspValTTPSerPheGlyValValLeuTyrglu 1014  
|||||  
|||||

Db 4455 TCTGATGTTTGAGCTTCGAGTGTCTCTATGAA 4490

RESULT 38  
US-09-844-353A-11

;; Sequence 11, Application US/09844353A  
;; Patent No. US20020037585A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ruvkun, Gary  
;; APPLICANT: Kimura, Gary  
;; APPLICANT: Patterson, Garth  
;; APPLICANT: Ogg, Scott  
;; APPLICANT: Paradis, Suzanne  
;; APPLICANT: Tissenbaum, Heidi  
;; APPLICANT: Morris, Jason  
;; APPLICANT: Kowek, Allison  
;; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR  
;; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS  
;; FILE REFERENCE: 00786/351005  
;; CURRENT APPLICATION NUMBER: US/09/844,353A  
;; CURRENT FILING DATE: 2001-04-27  
;; PRIOR APPLICATION NUMBER: US 08/857,076  
;; PRIOR FILING DATE: 1997-05-15  
;; NUMBER OF SEQ ID NOS: 114  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 11  
;; LENGTH: 5816  
;; TYPE: DNA  
;; ORGANISM: Caenorhabditis elegans  
;; FEATURE:  
;; NAME/KEY: misc.feature  
;; LOCATION: (1)...(5816)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-844-353A-11

Alignment Scores:  
Pred. No.: 0.0708 Length: 5816  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x US-09-844-353A-11 (1-5816)

OY 1303 SerAspValTTPSerPheGlyValValLeuTyrglu 1014  
|||||  
Db 4455 TCTGATGTTTGAGCTTCGAGTGTCTCTATGAA 4490

RESULT 39  
US-09-982-610-7

;; Sequence 7, Application US/09982610  
;; Patent No. US20020146420A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Genentech, Inc.  
;; APPLICANT: Bennett, Brian D.  
;; APPLICANT: Goeddel, David  
;; APPLICANT: Lee, James M.  
;; APPLICANT: Matthews, William  
;; APPLICANT: Tsai, Siao Ping  
;; APPLICANT: Wood, William I.  
;; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES  
;; NUMBER OF SEQUENCES: 45  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESS: Genentech, Inc.  
;; STREET: 460 Point San Bruno Blvd  
;; CITY: South San Francisco  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS



```
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;; FILE REFERENCE: 2077 001200
;; CURRENT APPLICATION NUMBER: US/09/796,692
;; CURRENT FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 8268
;; LENGTH: 510
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-796-692-8268

Alignment Scores:
Pred. No.: 0.887 Length: 510
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8268 (1-510)
Oy 940 CysValHisArgAspLeuAlaIaIaArgAsn 949
Db 375 TGTGTTCAAGAGACCTGCGCCGACGAGAC 404

RESULT 43
US-10-101-464A-355
;; Sequence 355, Application US/10101464A
;; Publication No. US20030046728A1
;; GENERAL INFORMATION:
;; APPLICANT: Strabala, Timothy
;; APPLICANT: Nieuwenhuizen, Nicolaas
;; APPLICANT: Higgins, Colleen M.
;; TITLE OF INVENTION: Compositions isolated from Plant Cells
;; FILE REFERENCE: 2077 001200
;; CURRENT APPLICATION NUMBER: US/10/101,464A
;; CURRENT FILING DATE: 2002-03-18
;; PRIOR APPLICATION NUMBER: 09/704,302
;; PRIOR FILING DATE: 2000-11-01
;; PRIOR APPLICATION NUMBER: 09/228,986
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/162,866
```

```
;; PRIOR FILING DATE: 1999-11-01
;; PRIOR APPLICATION NUMBER: PCT/US00/00724
;; PRIOR FILING DATE: 2000-01-11
;; NUMBER OF SEQ ID NOS: 989
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 355
;; LENGTH: 536
;; TYPE: DNA
;; ORGANISM: Pinus radiata
US-10-101-464A-355

Alignment Scores:
Pred. No.: 0.927 Length: 536
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-101-464A-355 (1-536)
Oy 959 LysIleAlaAspPheGlyLeuAlaIaLysIleu 968
Db 329 AAATATGCGGATTTGGATGCAAAACTT 358

RESULT 44
US-09-796-692-8255/C
;; Sequence 8255, Application US/09796692
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
;; FILE REFERENCE: 2077 001200
;; CURRENT APPLICATION NUMBER: US/09/796,692
;; CURRENT FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 8255
;; LENGTH: 539
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-796-692-8255

Alignment Scores:
Pred. No.: 0.932 Length: 539
Score: 10.00 Matches: 10
```



```
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 0.91%               Indels: 0
DB: 9                             Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8255 (1-539)
QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
   |||||||
Db 420 TGTGTCACAGAGACCTGGCCGCCAGAGAC 391

RESULT 45
US-09-796-692-7883
; Sequence 7883, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7883
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7883

Alignment Scores:
Pred. No.: 0.958      Length: 556
Score: 10.00          Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 0.91%      Indels: 0
DB: 9                Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-7883 (1-556)
QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
   |||||||
Db 418 TGTGTCACAGAGACCTGGCCGCCAGAGAC 447

RESULT 46
US-09-796-692-9025
; Sequence 9025, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9025
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-9025

Alignment Scores:
Pred. No.: 0.97      Length: 564
Score: 10.00          Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 0.91%      Indels: 0
DB: 9                Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-9025 (1-564)
QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
   |||||||
Db 418 TGTGTCACAGAGACCTGGCCGCCAGAGAC 447

RESULT 47
US-09-796-692-8630
; Sequence 8630, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8630
; LENGTH: 564
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8630

Alignment Scores:
Pred. No.: 0.97      Length: 564
Score: 10.00          Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 0.91%      Indels: 0
DB: 9                Gaps: 0
```

```
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 8630
;; LENGTH: 589
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (50)
;; OTHER INFORMATION: n-A,T,C or G
;; NAME/KEY: (231)
;; LOCATION: (231)
;; OTHER INFORMATION: n-A,T,C or G
US-09-796-692-8630
```

```
Alignment Scores:
Pred. No.: 1.01 Length: 589
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0
```

US-09-397-967-16 (1-1099) x US-09-796-692-8630 (1-589)

```
Oy 940 CysValHisArgSpleuAlaIlaArgasn 949
Db 418 TGTGTTCAAGAGACCTGGCCGCGAGAAC 447
```

RESULT 48

```
;; Sequence 7976, Application US/09796692
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
;; FILE REFERENCE: 2077.001200
;; CURRENT APPLICATION NUMBER: US/09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
```

```
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 7976
;; LENGTH: 597
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-796-692-7976
```

```
Alignment Scores:
Pred. No.: 1.02 Length: 597
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0
```

US-09-397-967-16 (1-1099) x US-09-796-692-7976 (1-597)

```
Oy 940 CysValHisArgSpleuAlaIlaArgasn 949
Db 418 TGTGTTCAAGAGACCTGGCCGCGAGAAC 447
```

RESULT 49

```
;; Sequence 9010, Application US/09796692
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
;; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
;; FILE REFERENCE: 2077.001200
;; CURRENT APPLICATION NUMBER: US/09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9010
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-9010

Alignment Scores:
Pred. No.: 1.02 Length: 597
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-9010 (1-597)

Oy 940 CysValHisArgAspLeuAlaAlaArgAsn 949
Db 418 TGTGTTCAAGAGACCTGGCCGCCAGGAAC 447

RESULT 50
US-09-796-692-8273
; Sequence 8273, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT FILING DATE: 2001-03-01
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8273
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (586)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-8273

Alignment Scores:
Pred. No.: 1.03 Length: 601
Score: 10.00 Matches: 10
```

```

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8273 (1-601)

Oy 940 CysValHisArgAspLeuAlaAlaArgAsn 949
Db 418 TGTGTTCAAGAGACCTGGCCGCCAGGAAC 447

RESULT 51
US-09-796-692-8047
; Sequence 8047, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT FILING DATE: 2001-03-01
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8047
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (542)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (592)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-8047

Alignment Scores:
Pred. No.: 1.03 Length: 605
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8047 (1-605)
```

```
Oy 940 CysValHisArgAspLeuAlaAlaArgAsn 949
|||||
Db 418 TGTGTTCAAGAGACCTGCGCCGACGAGAAC 447

RESULT 52
US-09-796-692-8927
; Sequence 8927, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Aigate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8927
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8927

Alignment Scores:
Pred. No.: 1.03 Length: 605
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8927 (1-605)
Oy 940 CysValHisArgAspLeuAlaAlaArgAsn 949
|||||
Db 418 TGTGTTCAAGAGACCTGCGCCGACGAGAAC 447

RESULT 53
US-09-796-692-8132/c
; Sequence 8132, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Aigate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8132
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8132

Alignment Scores:
Pred. No.: 1.04 Length: 608
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8132 (1-608)
Oy 940 CysValHisArgAspLeuAlaAlaArgAsn 949
|||||
Db 420 TGTGTTCAAGAGACCTGCGCCGACGAGAAC 391

RESULT 54
US-09-796-692-8085
; Sequence 8085, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Aigate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8132
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8132
```

```
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8132
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8132

Alignment Scores:
Pred. No.: 1.04 Length: 608
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-796-692-8132 (1-608)
Oy 940 CysValHisArgAspLeuAlaAlaArgAsn 949
|||||
Db 420 TGTGTTCAAGAGACCTGCGCCGACGAGAAC 391

RESULT 54
US-09-796-692-8085
; Sequence 8085, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Aigate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8132
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8132
```

```

; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 8085
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8085
```

```

Alignment Scores:
Pred. No.: 1.04 Length: 610
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0
```

US-09-397-967-16 (1-1099) x US-09-796-692-8085 (1-610)

```
Qy 940 CysValHisArgAspLeuAlaIArgAsn 949
Db 418 TGTGTCACAGAGACTGCGCCGACAGAAC 447
```

```

RESULT 55
US-09-796-692-7538
; Sequence 7538, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077, 001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 7538
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7538
```

```

Alignment Scores:
Pred. No.: 1.09 Length: 645
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0
```

US-09-397-967-16 (1-1099) x US-09-796-692-7538 (1-645)

```
Qy 940 CysValHisArgAspLeuAlaIArgAsn 949
Db 418 TGTGTCACAGAGACTGCGCCGACAGAAC 447
```

```

RESULT 56
US-10-101-464A-368
; Sequence 368, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000,1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1998-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 368
; LENGTH: 869
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-101-464A-368
```

US-09-397-967-16 (1-1099) x US-10-101-464A-368 (1-869)

```

Alignment Scores:
Pred. No.: 1.43 Length: 869
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0
```

US-09-397-967-16 (1-1099) x US-10-101-464A-368 (1-869)

```
Qy 959 LysIleAlaAspPheGlyLeuAlaLysLeu 968
Db 240 AAGATTGCAGATTGTTGCTGCGCTAAGCTT 269
```

```

RESULT 57
US-10-101-464A-218
; Sequence 218, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
```

```

; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 218
; LENGTH: 1021
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-218

Alignment Scores:
Pred. No.: 1.65      Length: 1021
Score: 10.00      Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.91%      Indels: 0
DB: 9      Gaps: 0

US-09-397-967-16 (1-1099) x US-10-101-464A-218 (1-1021)

QY 959 LysilealaaspheglyLeuAlaLysleu 968
Db 379 AAGATTGCGGATTTGGTCTGCGCAAGCTA 408

RESULT 58
US-10-101-464A-343
; Sequence 343, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 343
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-101-464A-343

Alignment Scores:
Pred. No.: 1.8      Length: 1128
Score: 10.00      Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.91%      Indels: 0
DB: 9      Gaps: 0

US-09-397-967-16 (1-1099) x US-10-101-464A-343 (1-1128)

```

---

```

QY 959 LysilealaaspheglyLeuAlaLysleu 968
Db 728 AAATAGCAGATTTTGGTTTACCCAACTC 757

RESULT 59
US-10-101-464A-872
; Sequence 872, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101.464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 872
; LENGTH: 2040
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-872

Alignment Scores:
Pred. No.: 3.07      Length: 2040
Score: 10.00      Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.91%      Indels: 0
DB: 9      Gaps: 0

US-09-397-967-16 (1-1099) x US-10-101-464A-872 (1-2040)

QY 959 LysilealaaspheglyLeuAlaLysleu 968
Db 1361 AAGATTGCGGATTTGGTCTGCGCAAGCTA 1390

RESULT 60
US-09-938-842A-1497
; Sequence 1497, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1497
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1497

```

Alignment Scores:

Pred. No.:	3 08	Length:	2052
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.91%	Indels:	0
DB:	9	Gaps:	0

US-09-397-967-16 (1-1099) x US-09-938-842A-1497 (1-2052)

QY 959 Lys11eAlaaspheg1yLeuAlaLysLeu 968  
|||||

DB 1465 AAATCCAGATTGATTGACTTAACCTC 1494

RESULT 61

US-09-938-842A-1480  
; Sequence 1480, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Mang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1480  
; LENGTH: 2136  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1480

Alignment Scores:

Pred. No.:	3 2	Length:	2136
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.91%	Indels:	0
DB:	9	Gaps:	0

US-09-397-967-16 (1-1099) x US-09-938-842A-1480 (1-2136)

QY 959 Lys11eAlaaspheg1yLeuAlaLysLeu 968  
|||||

DB 1585 AAATCCAGATTGATTGACTTAACCTC 1614

RESULT 62

US-09-924-859A-4  
; Sequence 4, Application US/09924859A  
; Patent No. US20020137113A1  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Mark, Melanie R.  
; APPLICANT: Sadick, Michael D.  
; APPLICANT: Shelton, David L.  
; APPLICANT: Wong, Wei Lee Tan  
; TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY  
; FILE REFERENCE: P0854C1P2C1  
; CURRENT APPLICATION NUMBER: US/09/924,859A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US/09/417,381  
; PRIOR FILING DATE: 1999-10-13  
; NUMBER OF SEQ ID NOS: 11

; SEQ ID NO 4  
; LENGTH: 2820  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-09-924-859A-4

Alignment Scores:

Pred. No.:	4 1	Length:	2820
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.91%	Indels:	0
DB:	10	Gaps:	0

US-09-397-967-16 (1-1099) x US-09-924-859A-4 (1-2820)

QY 1003 SeraspValTrpSerPheGlyValValLeu 1012  
|||||

DB 2468 AGCGACGTGTGAGACTTCGCGTGTCTC 2497

RESULT 63

US-09-982-610-22  
; Sequence 22, Application US/09982610  
; Patent No. US20020146420A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; Bennett, Brian D.  
; Goeddel, David  
; Lee, James M.  
; Matthews, William  
; Tsai, Siao Ping  
; Wood, William I.  
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE ACONIST ANTIBODIES  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Winpatin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/982,610  
; FILING DATE: 17-Oct-2001  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/446,648  
; FILING DATE: 1996-MAY-23  
; APPLICATION NUMBER: 08/222616  
; FILING DATE: 04-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0821P3PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1994  
; TELEFAX: 415/952-9881  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3120 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-982-610-22

Alignment Scores:

```
Pred. No.: 4.49 Length: 3120
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-982-610-22 (1-3120)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
|||||
Db 2419 TGTGTCCACAGACACTGCCCGCAGCAAC 2448

RESULT 64
US-09-866-510-1
; Sequence 1, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERN-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3270
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3267)
US-09-866-510-1

Alignment Scores:
Pred. No.: 4.68 Length: 3270
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-1 (1-3270)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
|||||
Db 2440 TGTGTCCACCGTGATCTGCTCTCGCAAC 2469

RESULT 65
US-09-866-510-3
; Sequence 3, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERN-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3270
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3267)
US-09-866-510-3

Alignment Scores:
Pred. No.: 4.68 Length: 3270
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-3 (1-3270)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
|||||
Db 2440 TGTGTCCACCGTGATCTGCTCTCGCAAC 2469

RESULT 66
US-09-866-510-5
; Sequence 5, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERN-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3270
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3267)
US-09-866-510-5

Alignment Scores:
Pred. No.: 4.68 Length: 3270
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-5 (1-3270)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949
|||||
Db 2440 TGTGTCCACCGTGATCTGCTCTCGCAAC 2469

RESULT 67
US-09-866-510-9
; Sequence 9, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERN-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
```



```

; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3270
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3267)
US-09-866-510-9

Alignment Scores:
Pred. No.: 4.68 Length: 3270
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-9 (1-3270)
QY 940 CysValHisArgAspLeuAlaAlaArgAsn 949
Db 2440 TGTGTCCACGCGTGTGCTGCTGCAGAC 2469

RESULT 68
US-09-866-510-13
; Sequence 13, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 3321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3270)
US-09-866-510-13

Alignment Scores:
Pred. No.: 4.74 Length: 3321
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-13 (1-3321)
QY 940 CysValHisArgAspLeuAlaAlaArgAsn 949
Db 2464 TCGGTCCACAGAGACTGCGGCTAGAGAC 2493

RESULT 69
US-09-866-510-15
; Sequence 15, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 3321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3318)
US-09-866-510-15

Alignment Scores:
Pred. No.: 4.74 Length: 3321
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-15 (1-3321)
QY 940 CysValHisArgAspLeuAlaAlaArgAsn 949
Db 2464 TCGGTCCACAGAGACTGCGGCTAGAGAC 2493

RESULT 70
US-09-866-510-17
; Sequence 17, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 3321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3318)
US-09-866-510-17

Alignment Scores:
Pred. No.: 4.74 Length: 3321
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-17 (1-3321)
QY 940 CysValHisArgAspLeuAlaAlaArgAsn 949
Db 2464 TCGGTCCACAGAGACTGCGGCTAGAGAC 2493
```

RESULT 71  
US-09-866-510-21  
; Sequence 21, Application US/09866510  
; Patent No. US20020111304A1  
; GENERAL INFORMATION:  
; APPLICANT: KAZLAUSKAS, ANDRIUS  
; APPLICANT: IKONO, YASUSHI  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES  
; FILE REFERENCE: ERM-104.01  
; CURRENT APPLICATION NUMBER: 2001-05-25  
; PRIORITY FILING DATE: 2001-05-25  
; PRIORITY FILING DATE: 2000-12-01  
; PRIORITY FILING DATE: 2000-12-01  
; PRIORITY FILING DATE: 2001-05-07  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 3321  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(3318)  
US-09-866-510-21

Alignment Scores:  
Pred. No.: 4.74 Length: 3321  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x US-09-866-510-21 (1-3321)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949  
|||||  
DB 2464 TGCCTCCACAGACCTGGCGCTAGGAC 2493

RESULT 72  
US-09-919-408-1  
; Sequence 1, Application US/09919408  
; Patent No. US20020072077A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Imclone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/919,408  
; FILING DATE: 31-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/977,451  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3453 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 112..3006  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 31..111  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31..3009  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-919-408-1

Alignment Scores:  
Pred. No.: 4.91 Length: 3453  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x US-09-919-408-1 (1-3453)

QY 940 CysValHisArgAspLeuAlaIaArgAsn 949  
|||||  
DB 2458 TGTGTCCACAGACCTGGCAGCCAGGAT 2487

RESULT 73  
US-09-872-136-1  
; Sequence 1, Application US/09872136  
; Patent No. US20020119545A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Imclone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/872,136  
FILING DATE: 01-Jun-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/208,786  
FILING DATE: <unknown>  
APPLICATION NUMBER: US/09/021,324  
FILING DATE: <unknown>  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 1992-11-19  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3453 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 112..3006  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 31..111  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31..3009  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-872-136-1  
Alignment Scores:  
Pctd. No.: 4.91 Length: 3453  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
DB: 10 Gaps: 0  
US-09-397-967-16 (1-1099) x US-09-872-136-1 (1-3453)  
QY 940 CysValHisArgAspLeuAlaIaArgAsn 949  
|||||  
DB 2458 TGTGTCCACAGACCTGCGCAGCGAAT 2487

RESULT 74  
US-09-919-408-3  
Sequence 3, Application US/09919408  
Patent No. US20020072077A1  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TROPICENT HEMATOPOIETIC STEM CELL  
RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Imclone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/919,408  
FILING DATE: 31-Jul-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/977,451  
FILING DATE: <unknown>  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3501 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 58..3039  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 139..3036  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 58..138  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-919-408-3

Alignment Scores:  
Pred. No.: 4.97 Length: 3501  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-919-408-3 (1-3501)

QY 940 CysValHisArgAspLeuAlaIArgAsn 949  
|||||  
Db 2476 TGTGTTCACAGACCTGGCCGACGAGAAC 2505

RESULT 75  
US-09-872-136-3  
Sequence 3, Application US/09872136  
Patent No. US20020119545A1

GENERAL INFORMATION:  
APPLICANT: Lemischka, Thor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Imclone Systems Incorporated  
STREET: 180 Varlick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/872,136  
FILING DATE: 01-Jun-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/208,786  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/09/021,324  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 1992-11-19  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3501 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal

FEATURE:  
NAME/KEY: CDS  
LOCATION: 58..3039

FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 139..3036

FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 58..138

SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-872-136-3

Alignment Scores:  
Pred. No.: 4.97 Length: 3501  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-872-136-3 (1-3501)

QY 940 CysValHisArgAspLeuAlaIArgAsn 949  
|||||  
Db 2476 TGTGTTCACAGACCTGGCCGACGAGAAC 2505

RESULT 76  
US-09-955-363-35  
Sequence 35, Application US/09955363  
Patent No. US20020173621A1

GENERAL INFORMATION:  
APPLICANT: Sledziewski Ph.D., Andrzej Z  
Bell, Lilijan A.  
Kindsvogel Ph.D., Wayne R.  
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS  
AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE  
FUSIONS

NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry  
STREET: 6300 Columbia Center  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98104-7092

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/955,363  
FILING DATE: 18-Sep-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/634,510  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/347,291  
FILING DATE: 02-MAY-1989

ATTORNEY/AGENT INFORMATION:  
NAME: Makl J.D., David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 990008.446C3

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-682-4900  
TELEFAX: 206-682-6031  
TELEX: 3723836

```

; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 4054 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: N
;   ANTI-SENSE: N
;   ORIGINAL SOURCE:
;     ORGANISM: Homo sapiens
;     IMMEDIATE SOURCE:
;     CLONE: p-alpha-17B
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 205..3471
;   OTHER INFORMATION:
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-955-363-35

Alignment Scores:
Pred. No.: 5 67 Length: 4054
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-955-363-35 (1-4054)
QY 940 CysvalHisArgAspLeuAlaAlaArgAsn 949
Db 2644 TGtGTCCACCGTGTGCTGCTGCCAC 2673

RESULT 77
US-10-022-939-1
; Sequence 1, Application US/10022939
; Publication No. US20030032160A1
; GENERAL INFORMATION:
;   APPLICANT: Kendall, Richard L.
;   APPLICANT: Thomas, Kenneth A.
;   APPLICANT: Mao, Xianzhi
;   APPLICANT: Tebben, Andrew
;   TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR
;   FILE REFERENCE: 19963YDB
;   CURRENT APPLICATION NUMBER: US/10/022,939
;   CURRENT FILING DATE: 2001-12-18
;   PRIOR APPLICATION NUMBER: 09/483,539
;   PRIOR FILING DATE: 2000-01-14
;   PRIOR APPLICATION NUMBER: 09/098,707
;   PRIOR FILING DATE: 1998-06-17
;   PRIOR APPLICATION NUMBER: 60/050,962
;   PRIOR FILING DATE: 1997-06-18
;   NUMBER OF SEQ ID NOS: 8
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 1
;   LENGTH: 4071
;   TYPE: DNA
;   ORGANISM: Human
;   US-10-022-939-1

Alignment Scores:
Pred. No.: 5 69 Length: 4071
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-022-939-1 (1-4071)
QY 942 HisArgAspLeuAlaAlaArgAsnIleLeu 951
Db 3076 CACAGGACCTGGCGCAGCAATATTCCTC 3105

RESULT 78
US-10-100-405A-1
; Sequence 1, Application US/10100405A
; Publication No. US20030055239A1
; GENERAL INFORMATION:
;   APPLICANT: Kendall, Richard L.
;   APPLICANT: Thomas, Kenneth A.
;   APPLICANT: Mao, Xianzhi
;   APPLICANT: Tebben, Andrew
;   TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR
;   FILE REFERENCE: 19963YDC
;   CURRENT APPLICATION NUMBER: US/10/100,405A
;   CURRENT FILING DATE: 2002-08-13
;   PRIOR APPLICATION NUMBER: 10/022,939
;   PRIOR FILING DATE: 2001-12-18
;   PRIOR APPLICATION NUMBER: 09/483,539
;   PRIOR FILING DATE: 2000-01-14
;   PRIOR APPLICATION NUMBER: 09/098,707
;   PRIOR FILING DATE: 1998-06-17
;   PRIOR APPLICATION NUMBER: 60/050,962
;   PRIOR FILING DATE: 1997-06-18
;   NUMBER OF SEQ ID NOS: 8
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 1
;   LENGTH: 4071
;   TYPE: DNA
;   ORGANISM: Human
;   US-10-100-405A-1

Alignment Scores:
Pred. No.: 5 69 Length: 4071
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-100-405A-1 (1-4071)
QY 942 HisArgAspLeuAlaAlaArgAsnIleLeu 951
Db 3076 CACAGGACCTGGCGCAGCAATATTCCTC 3105

RESULT 79
US-09-375-248-1
; Sequence 1, Application US/09375248
; Publication No. US20030026759A1
; GENERAL INFORMATION:
;   APPLICANT: Ferrell, Robert E.
;   APPLICANT: Altalo, Karl
;   APPLICANT: Finegold, David N.
;   APPLICANT: Karkkainen, Marika
;   TITLE OF INVENTION: SCREENING AND THERAPY FOR LYMPHATIC DISORDERS INVOLVING
;   FILE REFERENCE: 28967/35255A
;   CURRENT APPLICATION NUMBER: US/09/375,248
;   CURRENT FILING DATE: 1999-08-16
;   EARLIER APPLICATION NUMBER: PCT/US99/06133
;   EARLIER FILING DATE: 1999-03-26
;   NUMBER OF SEQ ID NOS: 28
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 1
;   LENGTH: 4111
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: (20)..(4111)
;     FEATURE:
;       OTHER INFORMATION: Human Flt4 (VEGFR-3) long form cDNA
;       US-09-375-248-1
```

## Alignment Scores:

Pred. No.: 5.74 Length: 4111  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-375-248-1 (1-4111)

QY 942 HisargaspLeuAlaAlaArgAsnIleLeu 951

Db 3122 CACAGAGACTGCTGCTCGGACATCTCTG 3151

## RESULT 80

US-09-982-610-31

Sequence 31, Application US/09982610

Patent No. US20020146420A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

Bennett, Brian D.

Goeddel, David

Lee, James M.

Matthews, William

Tsai, Siao Ping

Wood, William I.

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/982,610

FILING DATE: 17-Oct-2001

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/446,648

FILING DATE: 1996-MAY-23

APPLICATION NUMBER: 08/222616

FILING DATE: 04-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER: 40,378

REFERENCE/DOCKET NUMBER: P0821P3PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 31:

LENGTH: 4425 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-09-982-610-31

## Alignment Scores:

Pred. No.: 6.13 Length: 4425  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0

DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-982-610-31 (1-4425)

QY 942 HisargaspLeuAlaAlaArgAsnIleLeu 951

Db 3133 CACAGAGACTGCTGCTCGGACATCTCTG 3162

## RESULT 81

US-09-955-363-1

Sequence 1, Application US/09955363

Patent No. US20020173621A1

GENERAL INFORMATION:

APPLICANT: Sledziewski Ph.D., Andrzej Z

Bell, Lillian A.

Kindvogel Ph.D., Wayne R.

TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS

AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE

FUSIONS

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESS: Seed and Berry

STREET: 6300 Columbia Center

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/955,363

FILING DATE: 18-Sep-2001

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/634,510

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: US 07/347,291

FILING DATE: 02-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Makl J.D., David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 990008.446C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4465 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: N

ANTI-SENSE: N

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

DEVELOPMENTAL STAGE: Adult

TISSUE TYPE: Skin

CELL TYPE: fibroblasts

IMMEDIATE SOURCE:

CLONE: PR-1X1

FEATURE:

NAME/KEY: CDS

LOCATION: 354..3671

OTHER INFORMATION:

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-955-363-1

Alignment Scores:

```

Pred. No.: 6.18 Length: 4465
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-955-363-1 (1-4465)

QY 940 CysValHisArgAspLeuAlaIArgAsn 949
DB 2817 TGGCTCCACAGAGACCTGGCGCTAGGAC 2846

RESULT 82
US-09-870-759-119
; Sequence 119, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 119
; LENGTH: 4989
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)..(4149)
; OTHER INFORMATION:
US-09-870-759-119

Alignment Scores:
Pred. No.: 6.83 Length: 4989
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-870-759-119 (1-4989)

QY 1003 SerAspValTrpSerPheGlyValValLeu 1012
DB 3622 TCGGACGCTGTGCTTCGCGGCTGCTC 3651

RESULT 83
US-09-954-531-1383
; Sequence 1383, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents using Canc
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
```

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1383
; LENGTH: 5084
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-1383

Alignment Scores:
Pred. No.: 6.94 Length: 5084
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-954-531-1383 (1-5084)

QY 942 HisArgAspLeuAlaIArgAsnIleLeu 951
DB 2389 CACAGAGACTTGGCAGCCAGAAATATCTC 2418

RESULT 84
US-09-967-768A-277
; Sequence 277, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening using Sign
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 277
; LENGTH: 5084
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-277

Alignment Scores:
Pred. No.: 6.94 Length: 5084
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-967-768A-277 (1-5084)

QY 942 HisArgAspLeuAlaIArgAsnIleLeu 951
DB 2389 CACAGAGACTTGGCAGCCAGAAATATCTC 2418

RESULT 85
US-09-919-408-5
; Sequence 5, Application US/09919408
; Patent No. US20020072077A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Thor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
```

STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/919,408  
FILING DATE: 31-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/977,451  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401-  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5406 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
MOLECULE TYPE: linear  
TOPOLOGY: linear  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 208..4311  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 265..4308  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 208..264  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-919-408-5  
Alignment Scores:  
Pred. No.: 7.34 Length: 5406  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
DB: 10 Gaps: 0  
US-09-397-967-16 (1-1099) x US-09-919-408-5 (1-5406)  
QY 942 HisArgAspLeuAlaAlaArgAsnIleLeu 951  
DB 3277 CACAGGAGCCTGGCAGCAGAACATTTCTC 3306

RESULT 86  
US-09-872-136-5  
Sequence 5, Application US/09872136  
Patent No. US20020119545A1  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Imclone Systems Incorporated  
STREET: 180 Varlick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/872,136  
FILING DATE: 01-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/208,786  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/09/021,324  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 1992-11-19  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5406 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
MOLECULE TYPE: linear  
TOPOLOGY: linear  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 208..4311  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 265..4308  
FEATURE:



```

; NAME/KEY: sig.peptide
; LOCATION: 208..264
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-872-136-5

Alignment Scores:
Pred. No.: 7.34 Length: 5406
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-872-136-5 (1-5406)
Oy 942 HisArgAspLeuAlaAlaArgAsnIleLeu 951
Db 3277 CACAGGACCTGGCAGCAGCAACATTCTC 3306

RESULT 87
US-09-766-678-1
; Sequence 1, Application US/09766678
; Patent No. US20020081650A1
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; Rissau, Werner
; Millauer, Birgit
; Gazit, Aviv
; Levitzki, Alex
; TITLE OF INVENTION: Flk-1 Is A Receptor For Vascular
; Endothelial Growth Factor
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/766,678
; FILING DATE: 25-Jan-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/193,829
; FILING DATE: 09-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-060
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212)869-9741
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 286..4386
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-766-678-1

Alignment Scores:
Pred. No.: 7.41 Length: 5470
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-766-678-1 (1-5470)
Oy 942 HisArgAspLeuAlaAlaArgAsnIleLeu 951
Db 3355 CACAGGACCTGGCAGCAGCAACATTCTC 3384

RESULT 88
US-09-919-497-40
; Sequence 40, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 40
; LENGTH: 6378
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-497-40

Alignment Scores:
Pred. No.: 8.51 Length: 6378
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-919-497-40 (1-6378)
Oy 940 CysValHisArgAspLeuAlaAlaArgAsn 949
Db 2579 TGTGTCACCGATCTGCTGCTCGCAAC 2608

RESULT 89
US-09-769-987-1
; Sequence 1, Application US/09769987
; Patent No. US20020055129A1
; GENERAL INFORMATION:
; APPLICANT: Matsui, Toshimitsu
; Aaronson, Stuart A.
; APPLICANT: Pierce, Jacalyn H.
; TITLE OF INVENTION: Antibodies for the Alpha Platelet-Derived Growth Factor Recept
; FILE REFERENCE: 14014.026602
; CURRENT APPLICATION NUMBER: US/09/769,987
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 08/460,656
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: US 08/439,095
; PRIOR FILING DATE: 1995-05-11
; PRIOR APPLICATION NUMBER: US 07/915,884
; PRIOR FILING DATE: 1992-07-20
; PRIOR APPLICATION NUMBER: US 07/308,282
; PRIOR FILING DATE: 1989-02-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6412
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: No. US2002005129a1e =  
OTHER INFORMATION: synthetic construct  
NAME/KEY: CDS  
LOCATION: (139)...(3406)  
US-09-769-987-1

Alignment Scores:  
Pred. No.: 8.55 Length: 6412  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-769-987-1 (1-6412)

OY 940 CysValHisArgAspLeuAlaIArgAsn 949  
|||||  
Db 2578 TGTGTCCACCGTGTGCTGCTGCTGCAC 2607

RESULT 90  
US-09-982-610-17/c  
Sequence 17, Application US/09982610  
Patent No. US20020146420A1  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
Bennett, Brian D.  
Goeddel, David  
Lee, James M.  
Matthews, William  
Tsal, Siao Ping  
Wood, William I.  
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/982,610  
FILING DATE: 17-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/446,648  
FILING DATE: 1996-MAY-23  
APPLICATION NUMBER: 08/222616  
FILING DATE: 04-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0821P3PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6827 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-09-982-610-17

Alignment Scores:  
Pred. No.: 9.04 Length: 6827  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-982-610-17 (1-6827)

OY 942 HisArgAspLeuAlaIArgAsnIleu 951  
|||||  
Db 2671 CACAGAGACCTGCTGCTGGAACATTCTG 2642

RESULT 91  
US-09-982-610-45  
Sequence 45, Application US/09982610  
Patent No. US20020146420A1  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
Bennett, Brian D.  
Goeddel, David  
Lee, James M.  
Matthews, William  
Tsal, Siao Ping  
Wood, William I.  
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/982,610  
FILING DATE: 17-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/446,648  
FILING DATE: 1996-MAY-23  
APPLICATION NUMBER: 08/222616  
FILING DATE: 04-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0821P3PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9108 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 45:  
US-09-982-610-45

Alignment Scores:  
Pred. No.: 11.7 Length: 9108  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-982-610-45 (1-9108)

OY 942 HisArcspLeuAlaIarGsnIleu 951

DB 4065 CACAGGACCTGCTGCTGGACATTCTG 4094

## RESULT 92

US-09-960-352-10875/C

; Sequence 10875, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO. 10875

; LENGTH: 332

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 47-LIB188-004-Q1-E1-D4

US-09-960-352-10875

## Alignment Scores:

Pred. No.: 6.35 Length: 332

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.82% Indels: 0

DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-10875 (1-332)

OY 1002 GlnSerAspValTrpSerPheGlyVal 1010

DB 259 CACAGCGACGTGTGCTTTGGGGGTG 233

## RESULT 93

US-09-960-352-1177/C

; Sequence 1177, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO. 1177

; LENGTH: 355

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 06-LIB188-012-Q1-E1-B9

US-09-960-352-1177

## Alignment Scores:

Pred. No.: 6.74 Length: 355

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.82% Indels: 0

DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-1177 (1-355)

OY 1002 GlnSerAspValTrpSerPheGlyVal 1010

DB 259 CACAGCGACGTGTGCTTTGGGGGTG 233

## RESULT 94

US-10-101-464A-332

; Sequence 332, Application US/10101464A

; Publication No. US20030046728A1

; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Nicolaas

; APPLICANT: Higgins, Colleen M.

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; FILE REFERENCE: 11000.1020C2 and Their Use in the Modification of Plant Cell Signaling

; CURRENT APPLICATION NUMBER: US/10/101,464A

; CURRENT FILING DATE: 2002-03-18

; PRIOR APPLICATION NUMBER: 09/704,302

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 09/228,986

; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: 60/162,866

; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: PCT/US00/00724

; PRIOR FILING DATE: 2000-01-11

; NUMBER OF SEQ ID NOS: 989

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO. 332

; LENGTH: 363

; TYPE: DNA

; ORGANISM: Pinus radiata

US-10-101-464A-332

## Alignment Scores:

Pred. No.: 6.88 Length: 363

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.82% Indels: 0

DB: 9 Gaps: 0

US-09-397-967-16 (1-1099) x US-10-101-464A-332 (1-363)

OY 961 AlaAspPheGlyLeuAlaIarGsnIleu 969

DB 25 GCTGATTGCGCTTGCCAGCTGTG 51

## RESULT 95

US-09-960-352-10878

; Sequence 10878, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO. 10878

; LENGTH: 366

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 47-LIB188-007-Q1-E1-D4

US-09-960-352-10878

## Alignment Scores:

Pred. No.: 6.93 Length: 366

Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.82% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-10878 (1-366)

QY 1002 GlnSerAspValTrpSerPheGlyVal 1010  
|||||  
DB 289 CAGAGCGACGCTGTGCTTTGGGGTG 315

RESULT 96  
US-09-960-352-2603  
; Sequence 2603, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengding  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 2603  
; LENGTH: 391  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; OTHER INFORMATION: Clone ID: 12-LIB188-021-Q1-E1-C7

US-09-960-352-2603

Alignment Scores:  
Pred. No.: 7.35 Length: 391  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.82% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-2603 (1-391)

QY 1002 GlnSerAspValTrpSerPheGlyVal 1010  
|||||  
DB 291 CAGAGCGACGCTGTGCTTTGGGGTG 317

RESULT 97  
US-09-960-352-4293  
; Sequence 4293, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengding  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 4293  
; LENGTH: 392  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; OTHER INFORMATION: Clone ID: 19-LIB188-015-Q1-E1-E3

US-09-960-352-4293

Alignment Scores:  
Pred. No.: 7.37 Length: 392  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.82% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-4293 (1-392)

QY 1002 GlnSerAspValTrpSerPheGlyVal 1010  
|||||  
DB 291 CAGAGCGACGCTGTGCTTTGGGGTG 317

RESULT 98  
US-09-960-352-13765  
; Sequence 13765, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengding  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 13765  
; LENGTH: 401  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; OTHER INFORMATION: Clone ID: 59-LIB2809-008-Q1-E1-G12

US-09-960-352-13765

Alignment Scores:  
Pred. No.: 7.52 Length: 401  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.82% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-13765 (1-401)

QY 660 ProPheIleLysLeuSerAspProGly 668  
|||||  
DB 243 CCCTTCATCAGCTCAGTGACCTGGC 269

RESULT 99  
US-09-960-352-14442  
; Sequence 14442, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengding  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 14442  
; LENGTH: 402  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; OTHER INFORMATION: Clone ID: 62-LIB188-002-Q1-E1-H10

US-09-960-352-14442

Alignment Scores:  
Pred. No.: 7.54 Length: 402  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 0.82% Indels: 0  
 DB: 10 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-960-352-14442 (1-402)

QY 1002 GlnSerAspValTrpSerPheGlyVal 1010

DB 301 CAGAGCGACGTGTCTTTGGGGTG 327

RESULT 100

US-09-918-995-16302

; Sequence 16302, Application US/09918995  
 ; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 16302

; LENGTH: 415

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(415)

; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-16302

Alignment Scores:

Pred. No.:

7.75

Length:

415

Score:

9.00

Matches:

9

Percent Similarity:

100.00%

Conservative:

0

Best Local Similarity:

100.00%

Mismatches:

0

Query Match:

0.82%

Indels:

0

DB:

9

Gaps:

0

US-09-397-967-16 (1-1099) x US-09-918-995-16302 (1-415)

QY 941 ValHisArgAspLeuAlaAlaArgAsn 949

DB 133 GTACACAGGAGACTGGCCGCTCGAAGC 159

Search completed: April 28, 2003, 23:26:17  
 Job time : 451 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: April 28, 2003, 18:58:45 ; Search time 137 seconds  
(without alignments)  
2460.130 Million cell updates/sec

Title: US-09-397-967-16

Perfect score: 1099

Sequence: 1 MAPPERPLIPORSCSLSS.....RPAFATLSPOLDPLNMGRRG 1099

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Word size: 1

Total number of hits satisfying chosen parameters: 878600

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 150 summaries

Command line parameters:

-MODE=frame\_p2n.model -DEV=xlp  
-Q/cgn2.1/USPTO.spool/US09397967/runat.26042003.182927.11236/app.query.fasta.1.1287  
-DB=Issued Patents.NA -QFMT=fastapp -SUPER=oligo.rni -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=150 -DOCCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=100 -MODE=LOCAL  
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09397967@cgn2.1.49.@runat.26042003.182927.11236 -NCPU=6 -ICPU=3  
-NO\_XIPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-FEAPEXT=7 -TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

Issued Patents.NA:\*  
1: /cgn2.6/plodata/2/lna/5A.COMB.seq:\*  
2: /cgn2.6/plodata/2/lna/5B.COMB.seq:\*  
3: /cgn2.6/plodata/2/lna/5A.COMB.seq:\*  
4: /cgn2.6/plodata/2/lna/5B.COMB.seq:\*  
5: /cgn2.6/plodata/2/lna/PCTUS95.COMB.seq:\*  
6: /cgn2.6/plodata/2/lna/Backfillseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	16.4	4016	5	PCT-US95-08354A-1
2	62	5.6	3807	1	US-08-357-598-1
3	62	5.6	3807	2	US-09-003-289-1
4	62	5.6	3807	5	PCT-US95-16435-1
5	16	1.5	3435	4	US-09-046-158A-21
6	16	1.5	3495	1	US-08-446-038B-2
7	16	1.5	3495	1	US-08-446-010B-2
8	16	1.5	3495	1	US-08-805-445-2
9	16	1.5	3495	2	US-08-064-067D-2
10	16	1.5	3495	2	US-09-066-208-2
11	16	1.5	3495	4	US-08-980-080-3
12	16	1.5	3629	1	US-08-097-997A-8

13	16	1.5	3629	3	US-08-665-574C-8	Sequence 8, Appl1
14	16	1.5	3629	4	US-08-946-994-8	Sequence 8, Appl1
15	16	1.5	4482	2	US-08-567-508C-1	Sequence 1, Appl1
16	16	1.5	4482	3	US-09-196-480-1	Sequence 1, Appl1
17	13	1.2	41	1	US-08-481-003-9	Sequence 9, Appl1
18	13	1.2	41	3	US-08-485-598-9	Sequence 9, Appl1
19	13	1.2	45	1	US-08-481-003-8	Sequence 8, Appl1
20	13	1.2	45	3	US-08-485-598-8	Sequence 8, Appl1
21	12	1.1	151	1	US-08-222-616-11	Sequence 11, Appl1
22	12	1.1	151	5	US-08-446-642B-11	Sequence 11, Appl1
23	12	1.1	151	5	PCT-US95-04228-11	Sequence 11, Appl1
24	12	1.1	1398	2	US-08-604-989A-8	Sequence 8, Appl1
25	12	1.1	1398	2	US-08-604-989A-9	Sequence 9, Appl1
26	12	1.1	1521	2	US-08-604-989A-10	Sequence 10, Appl1
27	12	1.1	1713	2	US-09-741-154-1	Sequence 10, Appl1
28	12	1.1	1442	4	US-08-604-989A-11	Sequence 11, Appl1
29	12	1.1	1987	2	US-08-876-882-1	Sequence 11, Appl1
30	12	1.1	1987	2	US-08-876-882-1	Sequence 11, Appl1
31	12	1.1	2000	4	US-09-315-928-1	Sequence 1, Appl1
32	12	1.1	2000	4	US-08-426-509A-1	Sequence 1, Appl1
33	12	1.1	2000	5	PCT-US95-05008-1	Sequence 1, Appl1
34	12	1.1	2820	1	US-08-162-809-5	Sequence 1, Appl1
35	12	1.1	2862	2	US-08-449-645A-10	Sequence 10, Appl1
36	12	1.1	2862	2	US-08-702-367A-10	Sequence 10, Appl1
37	12	1.1	2962	5	PCT-US95-04681-10	Sequence 10, Appl1
38	12	1.1	2982	1	US-08-348-143-2	Sequence 2, Appl1
39	12	1.1	2982	1	US-08-571-785-2	Sequence 2, Appl1
40	12	1.1	2982	4	US-09-192-435-2	Sequence 2, Appl1
41	12	1.1	2982	4	US-09-558-340-2	Sequence 2, Appl1
42	12	1.1	3105	4	US-08-542-635-1	Sequence 2, Appl1
43	12	1.1	3116	2	US-08-449-645A-14	Sequence 14, Appl1
44	12	1.1	3116	2	US-08-702-367A-14	Sequence 14, Appl1
45	12	1.1	3116	5	PCT-US95-04681-14	Sequence 14, Appl1
46	12	1.1	3133	1	US-08-162-809-1	Sequence 1, Appl1
47	12	1.1	3150	4	US-09-166-350-23	Sequence 23, Appl1
48	12	1.1	3348	4	US-08-222-616-34	Sequence 34, Appl1
49	12	1.1	3348	4	US-08-446-648-34	Sequence 34, Appl1
50	12	1.1	3348	5	PCT-US95-04228-34	Sequence 34, Appl1
51	12	1.1	3546	1	US-08-162-809-9	Sequence 9, Appl1
52	12	1.1	3591	1	US-08-162-809-13	Sequence 13, Appl1
53	12	1.1	3752	2	US-08-466-537A-100	Sequence 100, App
54	12	1.1	3766	1	US-08-162-809-7	Sequence 7, Appl1
55	12	1.1	3969	1	US-08-436-044-5	Sequence 5, Appl1
56	12	1.1	3969	2	US-08-222-616-23	Sequence 23, Appl1
57	12	1.1	3969	2	US-08-436-054-5	Sequence 5, Appl1
58	12	1.1	3969	5	US-08-446-648-23	Sequence 23, Appl1
59	12	1.1	3969	5	PCT-US95-04228-23	Sequence 23, Appl1
60	12	1.1	4027	1	US-08-348-143-3	Sequence 3, Appl1
61	12	1.1	4027	1	US-08-348-143-3	Sequence 3, Appl1
62	12	1.1	4027	1	US-08-571-785-3	Sequence 3, Appl1
63	12	1.1	4027	1	US-08-571-785-4	Sequence 4, Appl1
64	12	1.1	4027	4	US-09-192-435-3	Sequence 3, Appl1
65	12	1.1	4027	4	US-09-192-435-4	Sequence 4, Appl1
66	12	1.1	4027	4	US-09-558-340-3	Sequence 3, Appl1
67	12	1.1	4027	4	US-09-558-340-4	Sequence 4, Appl1
68	12	1.1	4049	1	US-08-162-809-17	Sequence 17, Appl1
69	12	1.1	4097	1	US-08-162-809-11	Sequence 11, Appl1
70	12	1.1	4304	4	US-08-368-776A-1	Sequence 1, Appl1
71	12	1.1	4304	5	PCT-US96-00419-1	Sequence 1, Appl1
72	12	1.1	4508	5	PCT-US93-06251-34	Sequence 34, Appl1
73	12	1.1	4529	2	US-08-449-645A-16	Sequence 16, Appl1
74	12	1.1	4529	2	US-08-702-367A-16	Sequence 16, Appl1
75	12	1.1	4529	5	PCT-US95-04681-16	Sequence 16, Appl1
76	12	1.1	5816	4	US-08-857-076-11	Sequence 11, Appl1
77	12	1.1	16389	4	US-09-741-154-3	Sequence 3, Appl1
78	12	1.1	147	2	US-08-222-616-7	Sequence 7, Appl1
79	12	1.1	147	2	US-08-876-882-3	Sequence 3, Appl1
80	12	1.1	147	4	US-08-446-648-7	Sequence 7, Appl1
81	12	1.1	147	5	US-09-315-928-3	Sequence 3, Appl1
82	12	1.1	147	5	PCT-US95-04228-7	Sequence 7, Appl1
83	12	1.1	1590	1	US-08-278-089A-3	Sequence 3, Appl1
84	12	1.1	1601	1	US-07-934-393B-3	Sequence 3, Appl1
85	12	1.1	1601	2	US-08-838-957A-3	Sequence 3, Appl1

86	11	1.0	3056	1	US-08-162-809-21	Sequence 21, Appl
87	11	1.0	3059	1	US-08-162-809-3	Sequence 3, Appl
88	11	1.0	3125	1	US-08-162-809-19	Sequence 19, Appl
89	11	1.0	3132	1	US-08-167-919A-9	Sequence 9, Appl
90	11	1.0	3132	3	US-08-715-106-9	Sequence 9, Appl
91	11	1.0	3162	2	US-08-449-645A-12	Sequence 12, Appl
92	11	1.0	3162	2	US-08-702-367A-12	Sequence 12, Appl
93	11	1.0	3162	5	PCT-US95-04681-12	Sequence 12, Appl
94	11	1.0	3254	1	US-08-162-809-15	Sequence 15, Appl
95	11	1.0	3906	2	US-08-469-537A-102	Sequence 102, App
96	11	1.0	4138	1	US-08-323-474-1	Sequence 1, Appl
97	11	1.0	4138	5	PCT-US93-06093-1	Sequence 1, Appl
98	11	1.0	4165	1	US-08-442-248-1	Sequence 1, Appl
99	11	1.0	4165	1	US-08-440-815-1	Sequence 1, Appl
100	11	1.0	4165	4	US-08-486-449-1	Sequence 1, Appl
101	11	1.0	4175	1	US-07-934-333B-1	Sequence 1, Appl
102	11	1.0	4175	1	US-08-278-089A-1	Sequence 1, Appl
103	11	1.0	4175	2	US-08-838-957A-1	Sequence 1, Appl
104	11	1.0	4176	1	US-08-278-089A-5	Sequence 5, Appl
105	11	1.0	4176	2	US-08-838-957A-5	Sequence 5, Appl
106	11	1.0	4322	1	US-08-673-789-1	Sequence 1, Appl
107	10	0.9	584	5	PCT-US92-02750-5	Sequence 5, Appl
108	10	0.9	510	1	US-08-810-116-5	Sequence 5, Appl
109	10	0.9	510	2	US-07-930-548A-5	Sequence 5, Appl
110	10	0.9	1894	4	US-07-912-122-3	Sequence 3, Appl
111	10	0.9	1894	5	PCT-US93-06404-3	Sequence 3, Appl
112	10	0.9	2301	1	US-08-306-691B-23	Sequence 23, Appl
113	10	0.9	2301	5	PCT-US93-06251-78	Sequence 78, Appl
114	10	0.9	2820	1	US-08-286-305A-4	Sequence 4, Appl
115	10	0.9	2820	2	US-08-441-104A-4	Sequence 4, Appl
116	10	0.9	2820	4	US-08-440-816A-4	Sequence 4, Appl
117	10	0.9	2820	2	US-09-417-381A-4	Sequence 4, Appl
118	10	0.9	3120	4	US-08-222-616-22	Sequence 22, Appl
119	10	0.9	3120	5	PCT-US95-04228-22	Sequence 22, Appl
120	10	0.9	3120	5	US-07-813-593-1	Sequence 1, Appl
121	10	0.9	3453	1	US-07-977-451-1	Sequence 1, Appl
122	10	0.9	3453	1	US-07-946-507-1	Sequence 1, Appl
123	10	0.9	3453	1	US-08-252-517-1	Sequence 1, Appl
124	10	0.9	3453	1	US-07-906-397A-1	Sequence 1, Appl
125	10	0.9	3453	1	US-08-601-891-1	Sequence 1, Appl
126	10	0.9	3453	2	US-09-021-324-1	Sequence 1, Appl
127	10	0.9	3453	5	PCT-US92-02750-1	Sequence 1, Appl
128	10	0.9	3453	5	PCT-US92-05401-1	Sequence 1, Appl
129	10	0.9	3453	5	PCT-US92-09893-1	Sequence 1, Appl
130	10	0.9	3475	2	US-08-222-299-3	Sequence 3, Appl
131	10	0.9	3475	2	US-08-434-878-3	Sequence 3, Appl
132	10	0.9	3475	5	PCT-US95-03718-3	Sequence 3, Appl
133	10	0.9	3476	1	US-08-183-211-1	Sequence 1, Appl
134	10	0.9	3476	5	PCT-US95-00176A-1	Sequence 3, Appl
135	10	0.9	3501	1	US-07-977-451-3	Sequence 3, Appl
136	10	0.9	3501	1	US-08-252-517-3	Sequence 3, Appl
137	10	0.9	3501	1	US-07-906-397A-3	Sequence 3, Appl
138	10	0.9	3501	1	US-08-601-891-3	Sequence 3, Appl
139	10	0.9	3501	2	US-09-021-324-3	Sequence 3, Appl
140	10	0.9	3501	5	PCT-US92-05401-3	Sequence 3, Appl
141	10	0.9	3521	1	US-08-222-299-1	Sequence 1, Appl
142	10	0.9	3521	1	US-08-434-878-1	Sequence 1, Appl
143	10	0.9	3521	5	PCT-US95-03718-1	Sequence 1, Appl
144	10	0.9	3521	5	US-08-097-997A-12	Sequence 12, Appl
145	10	0.9	3561	1	US-08-665-574C-12	Sequence 12, Appl
146	10	0.9	3561	4	US-08-946-994-12	Sequence 12, Appl
147	10	0.9	3561	4	US-08-180-195-35	Sequence 35, Appl
148	10	0.9	4054	1	US-08-477-329-35	Sequence 35, Appl
149	10	0.9	4054	1		
150	10	0.9	4054	1		

## ALIGNMENTS

RESULT 1  
PCT-US95-08354A-1  
; Sequence 1, Application PC/TUS9508354A  
; GENERAL INFORMATION:

APPLICANT: Temple University - Of The	
APPLICANT: Commonwealth System of Higher Education	
TITLE OF INVENTION: JAK3 PROTEIN TYROSINE	
TITLE OF INVENTION: KINASE AND DNA ENCODING THE SAME	
NUMBER OF SEQUENCES: 9	
CORRESPONDENCE ADDRESS:	
ADDRESSEE: Seidel, Gonda, Layorgna	
ADDRESSEE: 6 Monaco, P.C.	
STREET: Suite 1800, Two Penn Center	
CITY: Philadelphia	
STATE: Pennsylvania	
COUNTRY: U.S.A.	
ZIP: 19102	
COMPUTER READABLE FORM:	
MEDIUM TYPE: Diskette, 3.50 Inch, 720 Kb	
COMPUTER: IBM PS/2	
OPERATING SYSTEM: MS-DOS	
SOFTWARE: WordPerfect 5.1	
CURRENT APPLICATION DATA:	
APPLICATION NUMBER: PCT/US95/08354A	
FILING DATE:	
CLASSIFICATION:	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: 08/272,368	
FILING DATE: 8 July 1994	
ATTORNEY/AGENT INFORMATION:	
NAME: Monaco, Daniel A.	
REGISTRATION NUMBER: 30,480	
REFERENCE/DOCKET NUMBER: 6056-203 PC	
TELEPHONE: (215) 568-8383	
TELEFAX: (215) 568-5549	
INFORMATION FOR SEQ ID NO: 1:	
SEQUENCE CHARACTERISTICS:	
LENGTH: 4016 nucleotides	
TYPE: nucleic acid	
STRANDEDNESS: single stranded	
TOPOLOGY: linear	
PCT-US95-08354A-1	
Alignment Scores:	
Pred. No.: 8.44e-154	Length: 4016
Score: 180.00	Matches: 529
Percent Similarity: 97.24%	Conservative: 0
Best Local Similarity: 97.24%	Mismatches: 9
Query Match: 16.38%	Indels: 15
Gaps: 0	
US-09-397-967-16 (1-1099) x PCT-US95-08354A-1 (1-4016)	
QY 535 ATGATGGUAlValAspSglYglUthrHsAspSerGIuValLeuLeuLysValMetasp 554	
DB 2257 AGCGGAGGAGTGTGATGAGTGAACATGACATGCTGGAAGTCTCTTAAGGTCATGGAC 2316	
QY 555 SerATGHSAsrAsrCysMetGluSerPheLeuGIuAlAlAserLeuMetSerGIuVal 574	
DB 2317 TTCAGNCATCGAATCGATGATGATCTTTTCGAAGCCGCAAGCTTGATGACCAAGA 2376	
QY 575 SerTYrProHIsLeuValLeuLeuHIsGlyValCysMetAlAGLysSerTleMetVal 594	
DB 2377 TCCTACCCGACCTGCTGTTACTGCACGCGCTCTCATGCTGGAGACAGCATCTGTTG 2436	
QY 595 GInGluPheValTYrLeuGlyAlAlLeaspMetTYrLeuArgLysArgLysHIsLeuVal 614	
DB 2437 CAGGAATTGTGTATCTAGACCAATTGACATGTACTGCGCAAGCTGGCCACTGTTG 2496	
QY 615 SerAlaserTYrPlysLeuGlnValThrLysGlnLeuAlATyAlaLeuAsnTYrLeuGlu 634	
DB 2497 TCAGCCAGCTGGAACACTGAGGTGACCAAGCAGCTGATATGCCCTTAACACTTGGAG 2556	
QY 635 AspLysGlyLeuProHIsGlyAsnValSerAlaArgLysValLeuLeuAlaArgLysGly 654	
DB 2557 GACAAAGGCTTCCTACGCAAGCAGCTCAGCAGCAAGAGTCTCTGCTGCTGATAGGGG 2616	



QY 655 G1AAspG1YasnProPheIleLysLeuSerAspProGlyValSerProThrValLeu 674  
 DB 2617 GGAGATGGGAAATCCACCTTTCAATTAAGCTAGATGATCGTGTGACGTCCACTGTGCTG 2676  
 QY 675 SerLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAla 694  
 DB 2677 AGCTGGAAATGCTCACACCGACAGAAATACCTGGGTGGCCCGCAATGTCTCCAGAGAGCT 2736  
 QY 695 GlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyValAlaThrThrTrpGluValPhe 714  
 DB 2737 CACACACTGCTGCTGGAGGTGACAAAGTGGGCTTTGGAGCCACAGTGGGAGGTGTT- 2795  
 QY 715 GlnArg-GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTrpGly 734  
 DB 2796 CACGGGGGGAGCCCGCCACATCCCTCGCTGGAGCCCGCCAAAACCTGAAGTCTATGA 2855  
 QY 734 uAspGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCys 754  
 DB 2856 GGACCAAGGAGACAGCTCCGCTCTCAATGACAGAACTGGCGGAGCTTATCACACAGTG 2915  
 QY 754 sMetAlaTrpAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuSngl 774  
 DB 2916 CATGGGCTATGATCTTGCCGGGCGCCCTCTCTCCAGCTATCTAGAGACTCAACGG 2975  
 QY 774 IleuIleThrSerAspTrpGlyLeuLeuSerAspProThrProGlyIleProSerProAr 794  
 DB 2976 CCTCATATCATCAGATTAGAGAGCTCTCCA-GACCCACACCTGGCATCCGAGTCTCTG 3034  
 QY 794 gAspGluLeuCysValAlaGlyValAlaGlnLeuTrpAlaCysGlnAspProAlaIlePheG1 814  
 DB 3035 AGATTGAGCTGTGG--T--GGGCCCCAGCTATAGCTTCCAGAGACCCCGCCCATATAC-GA 3089  
 QY 814 uGluArgHisLeuLysTrpIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLe 834  
 DB 3090 GGAGAGACACCTTAATCATCTTGTGTGGCAAGGCACTTTGGACAGCTGGAGCT 3149  
 QY 834 uCysArgTrpAspProLeuGlyAsnAsnThrGlyProLeuValAlaValLysGlnLeuG1 854  
 DB 3150 GTGCCCTGTGACCCCC--TG-GACAAATGCGGAGCCCTGCTGGCAGTGAACACGTACA 3206  
 QY 854 nHisSerVal-PROAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaL 874  
 DB 3207 GCACAGCCGG-GCCACAGCCAGAGAGGACTTCCAGCGGAGATTCCAGATCTCTTAAGGCTC 3265  
 QY 874 eHisSerAspPheIleValLysTrpArgGlyValSerTrpGlyProGlyArgGlnSerL 894  
 DB 3266 TCCACAGCGACTTCAATCGTCAGATACCGGGAGTCAAGTATGGCCAGAGTCCCAAGAGCC 3325  
 QY 894 eHisLeuValMetGluTrpGlyLeuProSerGlyCysLeuArgAspLeu-LeuGlnArgHis 913  
 DB 3326 TCCGGTGTGTGATGAGTACCTGCCACAGCGCTGCTGGAGAA-CTTCTCTGACAGCGCAT 3384  
 QY 914 --ArgGly-LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMet 932  
 DB 3385 CGCGGGGGCCCTGCACAGCCAGCGCTACTGCTGTCTGCGTCAATCTGCAAGGCGCATG 3444  
 QY 933 GluTrpLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
 DB 3445 GAGTACCTGGTGGCGCGGCTGCTACACCTGACCTGGCTGCCCGCAACATCTTGGTG 3504  
 QY 953 GluSerGluAlaHisValLysTrpIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGly 972  
 DB 3505 GAGACGAGAGGCTCATGTATAAATCCGGGACTTGGCTGCTGCAACCTGCTGCCCTGGGA 3564  
 QY 973 LysAspTrpTrpValValAlaArgGluProGlyGlnSerProIlePheTrpTrpAlaProGly 992  
 DB 3565 AAGAGCTACTAGCTGCTCGGAGAGCTGCCAAAGCCCATCTTTGGTATAGCCCCGGAG 3624  
 QY 993 SerLeuSerAspAsnIlePheSerArgIleSerAspValTrpSerPheGlyValValLeu 1012  
 DB 3625 TCCCTATCTGACAAATCTTCTCCGCAATCTGACGTGTGAGAGCTTGGAGTGTGTG 3684

QY 1013 TyrGluLeuPheThrTrpCysAspLysSerCysSerProSerAlaGluPheLeuArgMet 1032  
 DB 3685 TACGAGCTCTTCACTACCTACGCGACAGAGAGTGTGAGCCCATCCCTGATCTCCGCGCATG 3744  
 QY 1033 MetGlyProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGly 1052  
 DB 3745 ATGGGGCTGAGAGGTAAAGAGACCCCGCTCTGCGGCTCTGAGAGCTGTGAGAGAGGCG 3804  
 QY 1053 ArgArgLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCys 1072  
 DB 3805 CGAGCGCTCCACACACTCCACCTCCACCTGCCACAGAGAGTTCAGAGAGCTCATGACGTGTC 3864  
 RESULT 2  
 US-08-357-598-1  
 ? Sequence 1, Application US/08357598  
 ? Patent No. 5705625  
 ? GENERAL INFORMATION:  
 ? APPLICANT: CLVIN, CURT I.  
 ? APPLICANT: SMALL, DONALD  
 ? TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
 ? NUMBER OF SEQUENCES: 12  
 ? CORRESPONDENCE ADDRESS:  
 ? ADDRESS: Fish & Richardson P.C.  
 ? STREET: 4225 Executive Square, Suite 1400  
 ? CITY: La Jolla  
 ? STATE: CA  
 ? COUNTRY: USA  
 ? ZIP: 92037  
 ? COMPUTER READABLE FORM:  
 ? MEDIUM TYPE: Floppy disk  
 ? COMPUTER: IBM PC compatible  
 ? OPERATING SYSTEM: PC-DOS/MS-DOS  
 ? SOFTWARE: Patent Release #1.0, Version #1.30  
 ? CURRENT APPLICATION DATA: US/08/357, 598  
 ? APPLICATION NUMBER: US/08/357, 598  
 ? FILING DATE: 15-DEC-1994  
 ? CLASSIFICATION: 435  
 ? ATTORNEY/AGENT INFORMATION:  
 ? NAME: Haile, Lisa A.  
 ? REGISTRATION NUMBER: 38,347  
 ? REFERENCE/DOCKET NUMBER: 07265/033001  
 ? TELECOMMUNICATION INFORMATION:  
 ? TELEPHONE: 619/678-5070  
 ? TELEFAX: 619/678-5099  
 ? INFORMATION FOR SEQ ID NO: 1:  
 ? SEQUENCE CHARACTERISTICS:  
 ? LENGTH: 3807 base pairs  
 ? TYPE: nucleic acid  
 ? STRANDEDNESS: single  
 ? TOPOLOGY: linear  
 ? MOLECULE TYPE: DNA  
 ? US-08-357-598-1  
 Alignment Scores:  
 Pred. No.: 1,24e-46 Length: 3807  
 Score: 62.00 Matches: 96  
 Percent Similarity: 97.96% Conservative: 0  
 Best Local Similarity: 97.96% Mismatches: 1  
 Query Match: 5.64% Indels: 2  
 DB: 1 Gaps: 0  
 US-09-397-967-16 (1,1099) x US-08-357-598-1 (1-3807)  
 QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHis 957  
 DB 2965 CGCGGCTGCGTGACACCGGAGCTGCGCGCCGAAACATCTCTGTGGAGAGCGACACAC 3044  
 QY 958 ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTrpTrpVal 977  
 DB 3045 GTCAAGATCGCTGACTTGGGCTAGCTAGCTGTGCGGCTTGTGAC-AAAGACTTACTACT 3103  
 QY 977 lValArgGluProGlyGlnSerProIlePheTrpTrpAlaProGluSerLeuSerAspAs 997

Db 3104 GGTCCGGAGCCAGCCAGCCCAATTTCTGCTATGCCCGGCAATCCCTCTGGACAA 3163  
|||  
QY 997 nllpHeserArGlnserAspValTrrpSerPheGlyValValleuTyrGluLeuPheTh 1017  
|||  
Db 3164 CATTTCTCTCGCCAGTCAGACGCTGTGAGGTGGGGTGTCTCTGTACAGAGCTCTTAC 3223  
|||  
QY 1017 rTyrCyAspLysSerCysSerProSerAlaGluPheLeuArgMetGly 1034  
|||  
Db 3224 CTACTGGACAAAGCTGACAGCCCTCGCGGAGTCTCTCGGATGATGGGA 3275  
|||

## RESULT 3

US-09-003-289-1  
; Sequence 1, Application US/09003289  
; Patent No. 5916792  
; GENERAL INFORMATION:  
; APPLICANT: Civin, Curt I.  
; APPLICANT: Small, Donald  
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/003,289  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/357,598  
; FILING DATE: 15-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halle, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/033001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3807 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-09-003-289-1

## Alignment Scores:

Pred. No.: 1,24e-46 Length: 3807  
Score: 62.00 Matches: 96  
Percent Similarity: 97.96% Conservative: 0  
Best Local Similarity: 97.96% Mismatches: 1  
Query Match: 5.64% Indels: 2  
Gaps: 0

US-09-397-967-16 (1-1099) x US-09-003-289-1 (1-3807)

QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957  
|||  
Db 2985 CGCCCTGCGTGACCGACGACCTGGCCCGCAACATCTCTGTGAGAGCGAGGACAC 3044  
|||  
QY 938 ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977  
|||  
Db 3045 GTCAGAGTCGCTGAGCTTGGCTAGCTAGCTGCTGCCCTTGAC-AAAGACTACTACGT 3103  
|||  
QY 977 lValAlaArgLupProGlyGlnSerProIlePheTrrpYrAlaProGluSerLeuSerAspAs 997  
|||

Db 3104 GGTCCGGAGCCAGCCAGCCCAATTTCTGCTATGCCCGGCAATCCCTCTGGACAA 3163  
|||  
QY 997 nllpHeserArGlnserAspValTrrpSerPheGlyValValleuTyrGluLeuPheTh 1017  
|||  
Db 3164 CATTTCTCTCGCCAGTCAGACGCTGTGAGGTGGGGTGTCTCTGTACAGAGCTCTTAC 3223  
|||  
QY 1017 rTyrCyAspLysSerCysSerProSerAlaGluPheLeuArgMetGly 1034  
|||  
Db 3224 CTACTGGACAAAGCTGACAGCCCTCGCGGAGTCTCTCGGATGATGGGA 3275  
|||

## RESULT 4

PCT-US95-16435-1  
; Sequence 1, Application PC/TUS9516435  
; GENERAL INFORMATION:  
; APPLICANT: The Johns Hopkins University School of Medicine  
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/16435  
; FILING DATE: 15-DEC-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halle, Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/033W01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3807 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
PCT-US95-16435-1

## Alignment Scores:

Pred. No.: 1,24e-46 Length: 3807  
Score: 62.00 Matches: 96  
Percent Similarity: 97.96% Conservative: 0  
Best Local Similarity: 97.96% Mismatches: 1  
Query Match: 5.64% Indels: 2  
Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US95-16435-1 (1-3807)

QY 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis 957  
|||  
Db 2985 CGCCCTGCGTGACCGACGACCTGGCCCGCAACATCTCTGTGAGAGCGAGGACAC 3044  
|||  
QY 938 ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa 977  
|||  
Db 3045 GTCAGAGTCGCTGAGCTTGGCTAGCTAGCTGCTGCCCTTGAC-AAAGACTACTACGT 3103  
|||  
QY 977 lValAlaArgLupProGlyGlnSerProIlePheTrrpYrAlaProGluSerLeuSerAspAs 997  
|||  
Db 3104 GGTCCGGAGCCAGCCAGCCCAATTTCTGCTATGCCCGGCAATCCCTCTGGACAA 3163  
|||  
QY 997 nllpHeserArGlnserAspValTrrpSerPheGlyValValleuTyrGluLeuPheTh 1017  
|||

Db 3164 CATCTCTCTGCGAGCTGAGAGCTGTGGAGCTTGGGGTCTCTGTAACAGCTCTTAC 3223  
OY 1017 rTTCysAspIysSerGlyssSerProSerAlaGluPheLeuArgMetHetGly 1034  
Db 3224 CTACTGCGACAAAGCTGACAGCCCTCGGGCGAGTCTCTGGATGATGGGA 3275  
RESULT 5  
US-09-046-158A-21  
; Sequence 21, Application US/09046158A  
; Patent No. 6187552  
; GENERAL INFORMATION:  
; APPLICANT: Roberda, Steven L.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING INHIBITORS OF  
; TITLE OF INVENTION: JAK2/CYTOKINE RECEPTOR BINDING  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pharmacia & Upjohn Co., Intellectual Property  
; STREET: 301 Henrietta Street  
; CITY: Kalamazoo  
; STATE: MI  
; COUNTRY: USA  
; ZIP: 49001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/046,158A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Darnley Jr., James D.  
; REGISTRATION NUMBER: 33,673  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 616/833-2210  
; TELEFAX: 616/833-8897  
; TELEX: 224401  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3435 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-046-158A-21  
Alignment Scores:  
Pred. No.: 6.9e-05 Length: 3435  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 4 Gaps: 0  
US-09-397-967-16 (1-1099) x US-09-046-158A-21 (1-3435)  
OY 1003 SerAspValITrpSerPheGlyValIleuTYrGluLeuPheThTYr 1018  
Db 3106 TCAGATGTGTGGAGCTTGGAGTGTCTGTATGAACCTTTTCACATAC 3153  
RESULT 6  
US-08-446-038B-2  
; Sequence 2, Application US/08446038B  
; Patent No. 5658791  
; GENERAL INFORMATION:  
; APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;  
; APPLICANT: Harpur, Alisa  
; TITLE OF INVENTION: No. 5658791el Protein Tyrosine Kinase  
; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felte & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,038B  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/064,067  
; FILING DATE: 30-Jun-1993  
; APPLICATION NUMBER: PCT/US91/08889  
; FILING DATE: 26-No. 5658791-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: Australian PK3594/90  
; FILING DATE: 28-No. 5658791-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: Australian 88229/91  
; FILING DATE: 27-No. 5658791-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5658791man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5244  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-688-9200  
; TELEFAX: 212-638-3884  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3495 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: nucleic acid  
US-08-446-038B-2  
Alignment Scores:  
Pred. No.: 7e-05 Length: 3495  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 1 Gaps: 0  
US-09-397-967-16 (1-1099) x US-08-446-038B-2 (1-3495)  
OY 1003 SerAspValITrpSerPheGlyValIleuTYrGluLeuPheThTYr 1018  
Db 2695 TCAGATGTGTGGAGCTTGGAGTGTCTGTATGAACCTTTTCACATAC 2742  
RESULT 7  
US-08-446-010B-2  
; Sequence 2, Application US/08446010B  
; Patent No. 571618  
; GENERAL INFORMATION:  
; APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;  
; APPLICANT: Harpur, Alisa  
; TITLE OF INVENTION: No. 571618el Protein Tyrosine Kinase  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felte & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,010B  
FILING DATE: 19-May-1995  
CLASSIFICATION: 433  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/446,038  
FILING DATE: 19-May-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/064,067  
FILING DATE: 30-Jun-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-No. 5716818-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-No. 5716818-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian 88229/91  
FILING DATE: 27-No. 5716818-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Baer, Madeline F.  
REGISTRATION NUMBER: 36,437  
REFERENCE/DOCKET NUMBER: LUD 5244.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-838-3884  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
US-08-446-010B-2

Alignment Scores:  
Pred. No.: 7e-05 Length: 3495  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
Gaps: 0  
DB: 1

US-09-397-967-16 (1-1099) x US-08-446-010B-2 (1-3495)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrgluLeuPheThrTyr 1018  
|||||  
DB 2695 TCAGATGTGTGGAGCTTGAGTGTCTATACGAACTTTCACATAC 2742

RESULT 8  
US-08-805-445-2  
Sequence 2, Application US/08805445  
Patent No. 5821069  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;  
APPLICANT: Harpur, Ailsa  
TITLE OF INVENTION: No. 5821069e1 Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch,  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/805,445  
FILING DATE: 25-FEB-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/446,038  
FILING DATE: 19-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/064,067  
FILING DATE: 30-Jun-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-No. 5821069-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-No. 5821069-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian 88229/91  
FILING DATE: 27-No. 5821069-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5821069man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5244  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-838-3884  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
US-08-805-445-2

Alignment Scores:  
Pred. No.: 7e-05 Length: 3495  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
Gaps: 0  
DB: 1

US-09-397-967-16 (1-1099) x US-08-805-445-2 (1-3495)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrgluLeuPheThrTyr 1018  
|||||  
DB 2695 TCAGATGTGTGGAGCTTGAGTGTCTATACGAACTTTCACATAC 2742

RESULT 9  
US-08-064-067D-2  
Sequence 2, Application US/08064067D  
Patent No. 5852184  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;  
APPLICANT: Harpur, Ailsa  
TITLE OF INVENTION: No. 5852184e1 Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/064,067D  
FILING DATE: 30-Jun-1993

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US91/08889  
;; FILING DATE: 26-NOV-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: Australian PK3594/90  
;; FILING DATE: 28-NOV-1990  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: Australian 88229/91  
;; FILING DATE: 27-NOV-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Hanson, No. 5852184man D.  
;; REGISTRATION NUMBER: 30,946  
;; REFERENCE/DOCKET NUMBER: LUD 5244  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-688-9200  
;; TELEFAX: 212-838-3884  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 3495 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: nucleic acid  
US-08-064-067D-2  
  
Alignment Scores:  
Pred. No.: 7e-05 Length: 3495  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: Gaps: 2  
  
US-09-397-967-16 (1-1099) x US-08-064-067D-2 (1-3495)  
QY 1003 setAspValTrpSerPheGlyValIleuTyrgLuleuPhehThyr 1018  
Db 2695 TCAGATGTGTGAGCTTTGGAGTGTCTATACGAACTTTTCACATAC 2742  
  
RESULT 10  
US-09-066-208-2  
; Sequence 2, Application US/09066208  
; Patent No. 5910426  
; GENERAL INFORMATION:  
; APPLICANT: Wilks, Andrew F.; Ziemlecki, Andrew;  
; APPLICANT: Harpur, Alisa  
; TITLE OF INVENTION: No. 5910426el Protein Tyrosine Kinase  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felte & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/066,208  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/805,445  
; FILING DATE: 25-FEB-1997  
; APPLICATION NUMBER: US 08/446,038  
; FILING DATE: 19-MAY-1995  
; APPLICATION NUMBER: 08/064,067  
; FILING DATE: 30-Jun-1993  
; APPLICATION NUMBER: PCT/US91/08889  
; FILING DATE: 26-NOV-1991

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: Australian PK3594/90  
;; FILING DATE: 28-NOV-1990  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: Australian 88229/91  
;; FILING DATE: 27-NOV-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Hanson, No. 5910426man D.  
;; REGISTRATION NUMBER: 30,946  
;; REFERENCE/DOCKET NUMBER: LUD 5244  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-688-9200  
;; TELEFAX: 212-838-3884  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 3495 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: nucleic acid  
US-09-066-208-2  
  
Alignment Scores:  
Pred. No.: 7e-05 Length: 3495  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: Gaps: 2  
  
US-09-397-967-16 (1-1099) x US-09-066-208-2 (1-3495)  
QY 1003 setAspValTrpSerPheGlyValIleuTyrgLuleuPhehThyr 1018  
Db 2695 TCAGATGTGTGAGCTTTGGAGTGTCTATACGAACTTTTCACATAC 2742  
  
RESULT 11  
US-08-980-080-3  
; Sequence 3, Application US/08980080  
; Patent No. 6312941  
; GENERAL INFORMATION:  
; APPLICANT: CARTER-SU, CHRISTIN  
; APPLICANT: ROI, LIANG-YOU  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING  
; TITLE OF INVENTION: SIGNALING PATHWAY AGONISTS AND ANTAGONISTS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL, LLP  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/980,080  
; FILING DATE: 26-NOV-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARROLL, PETER G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: UM-03036  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3495 base pairs  
; TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2979  
US-08-980-080-3

Alignment Scores:  
Pred. No.: 7e-05  
Score: 16.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 1.46%  
DB: 4

Length: 3495  
Matches: 16  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-980-080-3 (1-3495)

QY 1003 SeraspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
|||||  
DB 2695 TCACATGCTGACCTTGGAGTGTCTATACGACCTTTACACATAC 2742

RESULT 12  
US-08-097-997A-8  
Sequence 8, Application US/08097997A  
Patent No. 5728536  
GENERAL INFORMATION:  
APPLICANT: Ihle, James N.  
APPLICANT: Silvenoinen, Olli  
APPLICANT: Withuhn, Bruce A.  
APPLICANT: Quelle, Frederick W.  
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,997A  
FILING DATE: 29-JULY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656,0370000/SLE/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3629 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 94..3480  
US-08-097-997A-8

Alignment Scores:  
Pred. No.: 7.25e-05  
Score: 16.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 1.46%  
DB: 0

Length: 3629  
Matches: 16  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

Query Match: 1.46%  
DB: 1  
Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-097-997A-8 (1-3629)

QY 1003 SeraspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
|||||  
DB 3196 TCACATGCTGACCTTGGAGTGTCTATACGACCTTTACACATAC 3243

RESULT 13  
US-08-665-574C-8  
Sequence 8, Application US/08665574C  
Patent No. 6136595  
GENERAL INFORMATION:  
APPLICANT: Ihle, James N.  
APPLICANT: Silvenoinen, Olli  
APPLICANT: Withuhn, Bruce A.  
TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,574C  
FILING DATE: 18-JUN-1996  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/282,012  
FILING DATE: 29-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/097,997  
FILING DATE: 29-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/118,968  
FILING DATE: 09-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bugalsky, Lawrence B.  
REGISTRATION NUMBER: 35,086  
REFERENCE/DOCKET NUMBER: 0656,0370002/SLE/LBB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3629 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 94..3480  
US-08-665-574C-8

Alignment Scores:  
Pred. No.: 7.25e-05  
Score: 16.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 1.46%  
DB: 3

Length: 3629  
Matches: 16  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-665-574C-8 (1-3629)

Oy 1003 SeraspValTrpSerPheGlyValValleuTyrgIuleuPheThrTyr 1018  
|||||  
Db 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGACTTTTCACATAC 3243

## RESULT 14

US-08-946-994-8

Sequence 8, Application US/08946994  
Patent No. 6210654

## GENERAL INFORMATION:

APPLICANT: Ihle, James N.

APPLICANT: Sliwomoinen, Ollie

APPLICANT: Witthuhn, Bruce A.

APPLICANT: Queller, Frederick W.

TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein &amp; Fox

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/946,994

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/665,574

FILING DATE: 18-JUN-1996

APPLICATION NUMBER: 08/282,012

FILING DATE: 29-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/097,997

FILING DATE: 29-JUL-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/118,868

FILING DATE: 09-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/GKT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2540

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 3629 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 94..3480

US-08-946-994-8

Alignment Scores:

Pred. No.: 7.25e-05

Score: 16.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.468

DB: 4

Gaps: 0

US-09-397-967-16 (1-1099) x US-08-946-994-8 (1-3629)

Oy 1003 SeraspValTrpSerPheGlyValValleuTyrgIuleuPheThrTyr 1018  
|||||

Db 3196 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGACTTTTCACATAC 3243

## RESULT 15

US-08-567-508C-1

Sequence 1, Application US/08567508C  
Patent No. 5914393

## GENERAL INFORMATION:

APPLICANT: Coleman, Roger

APPLICANT: Stuart, Susan G.

TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/567,508C

FILING DATE: 05-DEC-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-004905

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4482 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

IMMEDIATE SOURCE:

LIBRARY: Placenta

CLONE: 179527

US-08-567-508C-1

Alignment Scores:

Pred. No.: 8.78e-05

Score: 16.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.468

DB: 2

Gaps: 0

US-09-397-967-16 (1-1099) x US-08-567-508C-1 (1-4482)

Oy 1003 SeraspValTrpSerPheGlyValValleuTyrgIuleuPheThrTyr 1018  
|||||

Db 3442 TCAGATGTGTGAGCTTTGGAGTGTCTCTATACGACTTTTCACATAC 3489

## RESULT 16

US-09-196-480-1

Sequence 1, Application US/09196480  
Patent No. 6019966

## GENERAL INFORMATION:

APPLICANT: Coleman, Roger

APPLICANT: Stuart, Susan G.

TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
City: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/196,480  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/567,508  
FILING DATE: 05-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-00490S  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4482 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
LIBRARY: Flaccenta  
CLONE: 179527  
US-09-196-480-1  
Alignment Scores:  
Pred. No.: 8.78e-05 Length: 4482  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 3 Gaps: 0  
US-09-397-967-16 (1-1099) x US-09-196-480-1 (1-4482)  
Oy 1003 SeraspValTrpSerPheGlyValLeuTYRGlulLeuPheThrTyr 1018  
|||||  
Db 3442 TCAGATGTTTGAGCTTGGAGTGTGATGAACCTTTCACATAC 3489  
RESULT 17  
US-08-481-003-9/c  
Sequence 9, Application US/08481003  
GENERAL INFORMATION:  
APPLICANT: CAPON, DANIEL J  
APPLICANT: TIAN, HUAN  
APPLICANT: SMITH, DOUGLAS H  
APPLICANT: WINSLOW, GENINE A  
APPLICANT: SIEKEVITZ, MIRIAM  
TITLE OF INVENTION: CHIMERIC RECEPTORS FOR REGULATING  
TITLE OF INVENTION: CELLULAR PROLIFERATION AND EFFECTOR FUNCTION  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CELL GENESYS, INC.  
STREET: 322 LAKESIDE DRIVE  
CITY: FOSTER CITY  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,003  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/382,846  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: KRUPEN, KAREN I  
REGISTRATION NUMBER: 34,647  
REFERENCE/DOCKET NUMBER: CELL 17  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 349-7392  
TELEFAX: (415) 349-7392  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 41 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-481-003-9  
Alignment Scores:  
Pred. No.: 0.000652 Length: 41  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 1 Gaps: 0  
US-09-397-967-16 (1-1099) x US-08-481-003-9 (1-41)  
Oy 1 MetalProProserGluGluThrProleuIleProGln 13  
|||||  
Db 40 ATGGACCTCCAGTGAAGAGACACCTCTGATCCCTCAG 2  
RESULT 18  
US-08-485-598-9/c  
Sequence 9, Application US/08485598  
Patent No. 6077947  
GENERAL INFORMATION:  
APPLICANT: CAPON, DANIEL J  
APPLICANT: TIAN, HUAN  
APPLICANT: SMITH, DOUGLAS H  
APPLICANT: WINSLOW, GENINE A  
APPLICANT: SIEKEVITZ, MIRIAM  
TITLE OF INVENTION: CHIMERIC RECEPTORS FOR REGULATING  
TITLE OF INVENTION: CELLULAR PROLIFERATION AND EFFECTOR FUNCTION  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CELL GENESYS, INC.  
STREET: 322 LAKESIDE DRIVE  
CITY: FOSTER CITY  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,598  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/382,846  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:



NAME: KRUPEN, KAREN I  
REGISTRATION NUMBER: 34,647  
REFERENCE/DOCKET NUMBER: CELL 17  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 358-9600 x131  
TELEFAX: (415) 349-7392  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 41 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-485-598-9

Alignment Scores:  
Pred. No.: 0.000652 Length: 41  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 3 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-485-598-9 (1-41)

OY 1 MetAlapProSerGluGluThrProLeuIleProGln 13  
DB 40 ATGGCACCTCCAACTGAGAGACACCTGATCCCTCAG 2

RESULT 19  
US-08-481-003-8

Sequence 8, Application US/08481003  
Patent No. 5741899

GENERAL INFORMATION:

APPLICANT: CAPON, DANIEL J

APPLICANT: TIAN, HUAN

APPLICANT: SMITH, DOUGLAS H

APPLICANT: WINSLOW, GENINE A

APPLICANT: SIEKEVITZ, MIRIAM

TITLE OF INVENTION: CHIMERIC RECEPTORS FOR REGULATING

TITLE OF INVENTION: CELLULAR PROLIFERATION AND EFFECTOR FUNCTION

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: CELL GENESYS, INC.

STREET: 322 LAKESIDE DRIVE

CITY: FOSTER CITY

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/481,003

FILING DATE:

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/382,846

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: KRUPEN, KAREN I

REGISTRATION NUMBER: 34,647

REFERENCE/DOCKET NUMBER: CELL 17

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 358-9600 x131

TELEFAX: (415) 349-7392

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 45 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-481-003-8

Alignment Scores:  
Pred. No.: 0.00071 Length: 45  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-481-003-8 (1-45)

OY 1 MetAlapProSerGluGluThrProLeuIleProGln 13  
DB 6 ATGGCACCTCCAACTGAGAGACACCTGATCCCTCAG 44

RESULT 20  
US-08-485-598-8

Sequence 8, Application US/08485598  
Patent No. 6077947

GENERAL INFORMATION:

APPLICANT: CAPON, DANIEL J

APPLICANT: TIAN, HUAN

APPLICANT: SMITH, DOUGLAS H

APPLICANT: WINSLOW, GENINE A

APPLICANT: SIEKEVITZ, MIRIAM

TITLE OF INVENTION: CHIMERIC RECEPTORS FOR REGULATING

TITLE OF INVENTION: CELLULAR PROLIFERATION AND EFFECTOR FUNCTION

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: CELL GENESYS, INC.

STREET: 322 LAKESIDE DRIVE

CITY: FOSTER CITY

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94404

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,598

FILING DATE:

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/382,846

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: KRUPEN, KAREN I

REGISTRATION NUMBER: 34,647

REFERENCE/DOCKET NUMBER: CELL 17

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 358-9600 x131

TELEFAX: (415) 349-7392

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 45 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)  
US-08-485-598-8

Alignment Scores:  
Pred. No.: 0.00071 Length: 45  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.18% Indels: 0  
DB: 3 Gaps: 0



APPLICANT: Wood, William I.  
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Palin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04228  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/222616  
FILING DATE: 04-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Wendy M. Lee  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 821P3PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US95-04228-11

Alignment Scores:  
Pred. No.: 0 0173 Length: 151  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US95-04228-11 (1-151)  
QY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952  
Db 1 GTGCACAGGAGATCTCGCGCTCGAACAATCTGCTC 36

RESULT 24  
US-08-604-989A-8  
Sequence 8, Application US/08604989A  
Patent No. 5834208  
GENERAL INFORMATION:  
APPLICANT: Sakano, S.  
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/604,989A  
FILING DATE: February 23, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Charles E. Miller  
REGISTRATION NUMBER: 24,576  
REFERENCE/DOCKET NUMBER: 1920-026  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 738 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: human  
STRAIN: UT-7  
US-08-604-989A-8

Alignment Scores:  
Pred. No.: 0.0732 Length: 738  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-604-989A-8 (1-738)  
QY 941 VALHISARGASPLEUALAALARGASNIIEUVAL 952  
Db 349 GTGCACGCGAGCTGCGCGCCGCAACAATCTGCTC 384

RESULT 25  
US-08-604-989A-9  
Sequence 9, Application US/08604989A  
Patent No. 5834208  
GENERAL INFORMATION:  
APPLICANT: Sakano, S.  
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/604,989A  
FILING DATE: February 23, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Charles E. Miller  
REGISTRATION NUMBER: 24,576  
REFERENCE/DOCKET NUMBER: 1920-026  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1398 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: human  
STRAIN: UT-7  
US-08-604-989A-9

Alignment Scores:  
Pred. No.: 0.131 length: 1398  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-604-989A-9 (1-1398)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
|||||  
DB 922 GTGCACCGCGACTGCGCCGCCGCAACATCCTGCTC 957

RESULT 26  
US-08-604-989A-10  
Sequence 10, Application US/08604989A  
Patent No. 5834208

GENERAL INFORMATION:  
APPLICANT: Sakano, S.  
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/604,989A  
FILING DATE: February 23, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Charles E. Miller  
REGISTRATION NUMBER: 24,576  
REFERENCE/DOCKET NUMBER: 1920-026  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1521 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: human  
STRAIN: UT-7  
US-08-604-989A-10

Alignment Scores:  
Pred. No.: 0.141 length: 1521  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-604-989A-10 (1-1521)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
|||||  
DB 1045 GTGCACCGCGACTGCGCCGCCGCAACATCCTGCTC 1080

RESULT 27  
US-09-741-154-1  
Sequence 1, Application US/09741154  
Patent No. 6437110

GENERAL INFORMATION:  
APPLICANT: BEASLEY, Ellen M. et al  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL001061  
CURRENT APPLICATION NUMBER: US/09/741,154  
CURRENT FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO: 1  
LENGTH: 1713  
TYPE: DNA  
ORGANISM: Human  
US-09-741-154-1

Alignment Scores:  
Pred. No.: 0.157 length: 1713  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-741-154-1 (1-1713)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
|||||  
DB 991 GTGCACCGCGACTGCGCCGCCGCAACATCCTGCTC 1026

RESULT 28  
US-08-604-989A-11  
Sequence 11, Application US/08604989A  
Patent No. 5834208

GENERAL INFORMATION:  
APPLICANT: Sakano, S.  
TITLE OF INVENTION: No. 5834208el Tyrosine Kinase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/604,989A  
FILING DATE: February 23, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Charles E. Miller  
REGISTRATION NUMBER: 24,576  
REFERENCE/DOCKET NUMBER: 1920-026  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE

;; INFORMATION FOR SEQ ID NO: 11:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1942 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA to mRNA  
;; ORIGINAL SOURCE:  
;; ORGANISM: human  
;; STRAIN: UT-7  
US-08-604-989A-11

## Alignment Scores:

Pred. No.:	0.177	Length:	1942
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-09-397-967-16 (1-1099) x US-08-604-989A-11 (1-1942)

Qy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 1252 GTGCACCGCGACCTGCGCGCCGCAACATCTCTGTC 1287

## RESULT 29

US-08-876-882-1

Sequence 1, Application US/08876882

Patent No. 5981201

GENERAL INFORMATION:

APPLICANT: Avraham, Hava

TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT

TITLE OF INVENTION: OF BREAST CANCER

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith &amp; Reynolds P.C.

STREET: Two Militta Drive

CITY: Lexington

STATE: MA

COUNTRY: USA

ZIP: 02173-4799

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/876,882

FILING DATE: 16-JUN-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/035,228

ATTORNEY/AGENT INFORMATION:

NAME: Doreen, Hogle M

REGISTRATION NUMBER: 36,361

REFERENCE/DOCKET NUMBER: NEDH97-01PA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 781-861-6240

TELEFAX: 781-861-9540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1987 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-876-882-1

Alignment Scores:

Pred. No.: 0.18

Length: 1987

Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-09-397-967-16 (1-1099) x US-08-876-882-1 (1-1987)

Qy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 1307 GTGCACCGCGACCTGCGCGCCGCAACATCTCTGTC 1342

## RESULT 30

US-09-315-928-1

Sequence 1, Application US/09315928

Patent No. 6368796

GENERAL INFORMATION:

APPLICANT: Avraham, Hava

TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF

TITLE OF INVENTION: BREAST CANCER

FILE REFERENCE: NEDH97-01PAZ

CURRENT APPLICATION NUMBER: US/09/315,928

CURRENT FILING DATE: 1999-05-20

PRIOR APPLICATION NUMBER: US 08/876,882

PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: US 60/035,228

PRIOR FILING DATE: 1997-01-08

NUMBER OF SEQ ID NOS: 5

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1

LENGTH: 1987

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (263)...(1846)

US-09-315-928-1

1

Alignment Scores:

Pred. No.:	0.18	Length:	1987
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	4	Gaps:	0

US-09-397-967-16 (1-1099) x US-09-315-928-1 (1-1987)

Qy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 1307 GTGCACCGCGACCTGCGCGCCGCAACATCTCTGTC 1342

## RESULT 31

US-08-426-509A-1

Sequence 1, Application US/08426509A

Patent No. 6326469

GENERAL INFORMATION:

APPLICANT: Ullrich, Axel

APPLICANT: Gishizky, Mikhail

TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN

TITLE OF INVENTION: TYROSINE KINASES

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie &amp; Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York,

STATE: NY

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTED Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/426,509A  
FILING DATE: 21-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/232,545  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-0074-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2000 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
US-08-426-509A-1

Alignment Scores:  
Pred. No.: 0.181 Length: 2000  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-426-509A-1 (1-2000)

QY 941 ValHisArgAspLeuAlaIaArgAsnIleLeuVal 952  
Db 1302 GTGCACCGCGACTGCGCCGCCGACACATCTGTGTC 1337

RESULT 32  
PCT-US95-05008-1  
Sequence 1, Application PC/TUS9505008  
GENERAL INFORMATION:  
APPLICANT: Sugen, Inc.  
APPLICANT: 515 Galveston Drive  
APPLICANT: Redwood City, California 94063-4720  
APPLICANT: United States of America  
APPLICANT: Wissenschaften E.V.  
APPLICANT: Hofgarten Str. 2  
APPLICANT: Munchen 80539  
APPLICANT: Germany  
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine  
TITLE OF INVENTION: Kinases  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05008  
FILING DATE: 24-APR-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/232,545

FILING DATE: 22-APR-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-0074  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)790-9090  
TELEFAX: (212)869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2000 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
PCT-US95-05008-1

Alignment Scores:  
Pred. No.: 0.181 Length: 2000  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US95-05008-1 (1-2000)

QY 941 ValHisArgAspLeuAlaIaArgAsnIleLeuVal 952  
Db 1302 GTGCACCGCGACTGCGCCGCCGACACATCTGTGTC 1337

RESULT 33  
US-08-162-809-5  
Sequence 5, Application US/08162809  
Patent No. 5457048  
GENERAL INFORMATION:  
APPLICANT: Pasquale, Elena B.  
APPLICANT: Sajjadi, Pereydown G.  
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CAMPBELL AND FLORES  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States of America  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/162,809  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 9503  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2820 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
FEATURE:

```

; NAME/KEY: CDS
; LOCATION: 2..2548
; US-08-162-809-5

Alignment Scores:
Pred. No.: 0.248 Length: 2820
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-5 (1-2820)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 1817 GTGCATCGGAGTCTAGCTGCTCGAACAATCTGCTC 1852

RESULT 34
US-08-449-645A-10
; Sequence 10, Application US/08449645A
; Patent No. 5981245
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,645A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2962 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2913
; US-08-449-645A-10

Alignment Scores:
Pred. No.: 0.259 Length: 2962
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-449-645A-10 (1-2962)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2179 GTTCACCGTGACCTGGCTGCCCGCAACATCTCTGTC 2214

RESULT 35
US-08-702-367A-10
; Sequence 10, Application US/08702367A
; Patent No. 5981246
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,367A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2962 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2913
; US-08-702-367A-10

Alignment Scores:
Pred. No.: 0.259 Length: 2962
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-702-367A-10 (1-2962)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2179 GTTCACCGTGACCTGGCTGCCCGCAACATCTCTGTC 2214

RESULT 36
PCT-US95-04681-10
; Sequence 10, Application PC/TUS9504681
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04681
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2962 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2913
PCT-US95-04681-10

Alignment Scores:
Pred. No.: 0.259 Length: 2962
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 5

US-09-397-967-16 (1-1099) x PCT-US95-04681-10 (1-2962)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2179 GTTCACCGTGACCTGCGTCCGACACATCCTGTC 2214

RESULT 37
US-08-348-143-2
Sequence 2, Application US/08348143
Patent No. 5506205
GENERAL INFORMATION:
APPLICANT: TAJIMA, HISAO
APPLICANT: KITAGAWA, KOICHIRO
APPLICANT: OHNO, HIROYUKI
APPLICANT: OHNO, TOSHIO
TITLE OF INVENTION: A No. 5506205el Polypeptide of protein p140 and DNAs
TITLE OF INVENTION: encoding it
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUCHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/348,143
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 315806/1993
FILING DATE: 24-NOV-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2982 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

```

```

TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: rat
TISSUE TYPE: skeletal muscle myoblast
CELL LINE: L6
US-08-348-143-2

Alignment Scores:
Pred. No.: 0.261 Length: 2982
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 1

US-09-397-967-16 (1-1099) x US-08-348-143-2 (1-2982)
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2248 GTGACCGTGACCTGCGTCCGACACATCCTGTC 2283

RESULT 38
US-08-571-785-2
Sequence 2, Application US/08571785
Patent No. 5804411
GENERAL INFORMATION:
APPLICANT: TAJIMA, HISAO
APPLICANT: KITAGAWA, KOICHIRO
APPLICANT: OHNO, HIROYUKI
APPLICANT: OHNO, TOSHIO
TITLE OF INVENTION: A No. 5804411el Polypeptide of protein p140
TITLE OF INVENTION: and DNAs encoding it
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUCHRUE, MION, ZINN, MACPEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/571,785
FILING DATE: 13-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/348,143
FILING DATE: 23-NOV-1994
APPLICATION NUMBER: JP 315806/1993
FILING DATE: 24-NOV-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2982 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: rat
TISSUE TYPE: skeletal muscle myoblast
CELL LINE: L6
US-08-571-785-2

Alignment Scores:

```



Pred. No.: 0.261 Length: 2982  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-571-785-2 (1-2982)

Qy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 2248 GTGCACCGTGACCTGCTGCCGCCGCAACATCTTGTG 2283

RESULT 39

US-09-192-435-2

Sequence 2, Application US/09192435

Patent No. 6303320

GENERAL INFORMATION:

APPLICANT: TAJIMA, HISAO

APPLICANT: KITAGAWA, KOICHIRO

APPLICANT: OHNO, HIROYUKI

APPLICANT: UENO, TOSHIO

TITLE OF INVENTION: A No. 6303320el Polypeptide of Protein p140

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESSES:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/192,435

FILING DATE:

CLASSIFICATION:

Prior Application DATA:

APPLICATION NUMBER: 08/571,785

FILING DATE: 13-DEC-1995

APPLICATION NUMBER: 08/348,143

FILING DATE: 23-NOV-1994

APPLICATION NUMBER: JP 315806/1993

FILING DATE: 24-NOV-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2982 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: rat

TISSUE TYPE: skeletal muscle myoblast

CELL LINE: L6

US-09-192-435-2

Alignment Scores:  
Pred. No.: 0.261 Length: 2982  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-192-435-2 (1-2982)

Qy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 2248 GTGCACCGTGACCTGCTGCCGCCGCAACATCTTGTG 2283

RESULT 40

US-09-558-340-2

Sequence 2, Application US/09558340

Patent No. 6432913

GENERAL INFORMATION:

APPLICANT: TAJIMA, HISAO

APPLICANT: KITAGAWA, KOICHIRO

APPLICANT: OHNO, HIROYUKI

APPLICANT: UENO, TOSHIO

TITLE OF INVENTION: A No. 6432913el Polypeptide of Protein p140

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESSES:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/558,340

FILING DATE: 26-APR-2000

CLASSIFICATION:

Prior Application DATA:

APPLICATION NUMBER: 09/192,435

FILING DATE: 08-JAN-1998

APPLICATION NUMBER: 08/571,785

FILING DATE: 13-DEC-1995

APPLICATION NUMBER: 08/348,143

FILING DATE: 23-NOV-1994

APPLICATION NUMBER: JP 315806/1993

FILING DATE: 24-NOV-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2982 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: rat

TISSUE TYPE: skeletal muscle myoblast

CELL LINE: L6

US-09-558-340-2

Alignment Scores:  
Pred. No.: 0.261 Length: 2982  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-558-340-2 (1-2982)

Qy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 2248 GTGCACCGTGACCTGCTGCCGCCGCAACATCTTGTG 2283

```

APPLICANT: Fox, Gary M.
TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,645A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3116 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 34..2994
US-08-449-645A-14

Alignment Scores:
Pred. No.: 0.271 Length: 3116
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB: 2

US-09-397-967-16 (1-1099) x US-08-449-645A-14 (1-3116)
QY 941 VALHISARGASPLEUALAIALARGASNIILELVAL 952
|||||
DB 2260 GTGATCTGTGATCTGGCGGCGAGACATCTGTGTG 2295

RESULT 43
US-08-702-367A-14
; Sequence 14, Application US/08702367A
; Patent No. 5961246
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,367A
FILING DATE:

```

```

: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Winter, Robert B.
: REFERENCE/DOCKET NUMBER: A-287
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3116 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 34..2994
: US-08-702-367A-14

Alignment Scores:
Pred. No.: 0.271 Length: 3116
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-702-367A-14 (1-3116)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2260 GTGCATCGTATCTGCGCCAGCAACATCTGCTG 2295

RESULT 44
: Sequence 14, Application PC/TUS9504681
: GENERAL INFORMATION:
: APPLICANT: Fox, Gary M.
: TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
: NUMBER OF SEQUENCES: 28
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Amgen Patent Operations/RBW
: STREET: 1840 Dehaven Drive
: CITY: Thousand Oaks
: STATE: California
: COUNTRY: USA
: ZIP: 91320
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/04681
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Winter, Robert B.
: REFERENCE/DOCKET NUMBER: A-287
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3116 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 34..2994
: PCT-US95-04681-14

Alignment Scores:
Pred. No.: 0.271 Length: 3116
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-1 (1-3133)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2125 GTCCACAGGATCTGCGCCAGCAACATCTGCTG 2160

RESULT 46
: Sequence 23, Application US/09166350A
: Patent No. 6440663
: GENERAL INFORMATION:
: APPLICANT: Pasquale, Elena B.
: TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CAMPBELL AND FLORES
: STREET: 4370 La Jolla Village Drive, Suite 700
: CITY: San Diego
: STATE: California
: COUNTRY: United States of America
: ZIP: 92122
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/162,809
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Campbell, Cathryn A.
: REGISTRATION NUMBER: 31,815
: REFERENCE/DOCKET NUMBER: P-LJ 9503
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-9001
: TELEFAX: (619) 535-8949
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3133 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: join(3..419, 421..2858)
: US-08-162-809-1

Alignment Scores:
Pred. No.: 0.273 Length: 3133
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 1 Gaps: 0
```

APPLICANT: Scanlan, Mathew  
APPLICANT: Chen, Yao  
APPLICANT: Stockert, Elisabeth  
APPLICANT: Old, Lloyd  
APPLICANT: Jager, Elke  
APPLICANT: Knuth, Alex  
TITLE OF INVENTION: Renal Cancer Associated Antigens and  
FILE REFERENCE: 10461/7051  
CURRENT APPLICATION NUMBER: US/09/166,350A  
EARLIER FILING DATE: 1998-10-05  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO: 23  
LENGTH: 3150  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-166-350-23

Alignment Scores:  
Pred. No.: 0.274 Length: 3150  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-166-350-23 (1-3150)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
DB 2236 GTTCACCGTGACCTGCTGCCGACATCTCTGTC 2271

RESULT 47

US-08-222-616-34  
Sequence 34, Application US/08222616

PATENT NO. 5635177  
GENERAL INFORMATION:  
APPLICANT: Bennett, Brian D.  
APPLICANT: Goeddel, David  
APPLICANT: Lee, James M.  
APPLICANT: Matthews, William  
APPLICANT: Tsai, Siao Ping  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST  
TITLE OF INVENTION: ANTIBODIES  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222,616  
FILING DATE: 4-APR-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00586  
FILING DATE: 22-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/826935  
FILING DATE: 22-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.

REGISTRATION NUMBER:  
REFERENCE/DOCKET NUMBER: 821P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3348 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-222-616-34

Alignment Scores:  
Pred. No.: 0.29 Length: 3348  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-222-616-34 (1-3348)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
DB 2227 GTGCATCTGTATCTGCCGACGACATCTCTGTC 2262

RESULT 48

US-08-446-648-34  
Sequence 34, Application US/08446648

PATENT NO. 6331302  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Bennett, Brian D.  
APPLICANT: Goeddel, David  
APPLICANT: Lee, James M.  
APPLICANT: Matthews, William  
APPLICANT: Tsai, Siao Ping  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST  
TITLE OF INVENTION: ANTIBODIES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,648  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/222616  
FILING DATE: 04-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0821P3PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3348 base pairs  
TYPE: Nucleic Acid

```

STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-446-648-34

Alignment Scores:
Pred. No.: 0 29
Score: 12.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.09%
DB: 4
Length: 3348
Matches: 12
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-446-648-34 (1-3348)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2227 GTGCATCGTATCTGGCCGACGAGACATCTCTGCTG 2262

RESULT 49
PCT-US95-04228-34
Sequence 34, Application PC/TUS9504228
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04228
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 3348 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-04228-34

Alignment Scores:
Pred. No.: 0 29
Score: 12.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.09%
DB: 5
Length: 3348
Matches: 12
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

```

```

US-09-397-967-16 (1-1099) x PCT-US95-04228-34 (1-3348)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2227 GTGCATCGTATCTGGCCGACGAGACATCTCTGCTG 2262

RESULT 50
US-08-162-809-9
Sequence 9, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
APPLICANT: Pasquale, Elena B.
APPLICANT: Sajjadi, Fereydon G.
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162,809
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9901
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 3546 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 2..2920
US-08-162-809-9

Alignment Scores:
Pred. No.: 0 305
Score: 12.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.09%
DB: 1
Length: 3546
Matches: 12
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-9 (1-3546)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2189 GTGCACCGAGACCTGGCTGCCGACACATCTCTGCTC 2224

RESULT 51
US-08-162-809-13
Sequence 13, Application US/08162809
Patent No. 5457048
GENERAL INFORMATION:
APPLICANT: Pasquale, Elena B.
APPLICANT: Sajjadi, Fereydon G.
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,

```

TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CAMPBELL AND FLORES  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States of America  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/162,809  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 9503  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3591 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..2965  
US-08-162-809-13  
Alignment Scores:  
Pred. No.: 0.309 Length: 3591  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0  
DB: 1  
US-09-397-967-16 (1-1099) x US-08-162-809-13 (1-3591)  
QY 941 ValHisArgAspLeuAlaIaIaArgAsnIleLeuVal 952  
Db 2234 GTGCACGAGACCTGGCTGCCCGCACATCCTGGTC 2269  
RESULT 52  
US-08-469-537A-100  
Sequence 100, Application US/08469537A  
GENERAL INFORMATION:  
APPLICANT: Maisongier, et al.  
TITLE OF INVENTION: EHK AND FOR TYROSINE  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: NY  
COUNTRY: U.S.A.  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,537A

FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/406,247  
FILING DATE: 17-MAR-1995  
APPLICATION NUMBER: USSN 08/144,992  
FILING DATE: 28-OCT-1993  
APPLICATION NUMBER: USSN 07/736,559  
FILING DATE: 26-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempler, Ph.D., Gall M  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: REG 070C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
TELEX:  
INFORMATION FOR SEQ ID NO: 100:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3592 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 598..3444  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: 56  
OTHER INFORMATION: /mod\_base= OTHER  
OTHER INFORMATION: /label= N  
OTHER INFORMATION: /note= "where N = G, A, C or T"  
FEATURE:  
NAME/KEY: modified\_base  
LOCATION: 3538  
OTHER INFORMATION: /mod\_base= OTHER  
OTHER INFORMATION: /label= N  
OTHER INFORMATION: /note= "where N = G, A, C or T"  
US-08-469-537A-100  
Alignment Scores:  
Pred. No.: 0.309 Length: 3592  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0  
DB: 2  
US-09-397-967-16 (1-1099) x US-08-469-537A-100 (1-3592)  
QY 941 ValHisArgAspLeuAlaIaIaArgAsnIleLeuVal 952  
Db 2977 GTTCATCGAGACCTAGACCTAGACATCTTGCG 3012  
RESULT 53  
US-08-162-809-7  
Sequence 7, Application US/08162809  
GENERAL INFORMATION:  
APPLICANT: Pasquale, Elena B.  
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,  
AND METHODS OF USE  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CAMPBELL AND FLORES  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States of America  
ZIP: 92122  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: #1.0, Version #1.25  
APPLICATION NUMBER: US/08/162,809  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 9503  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3776 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 290..3208  
US-08-162-809-7

Alignment Scores:  
Pred. No.: 0.323 Length: 3776  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-7 (1-3776)  
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
Db 2477 GTGCATCGTGTCTCGACGACGTACATCTTAGTC 2512

RESULT 54  
US-08-436-044-5  
Sequence 5, Application US/08436044  
Patent No. 5624899  
GENERAL INFORMATION:  
APPLICANT: Bennett, Brian D.  
APPLICANT: Matthews, William  
TITLE OF INVENTION: HTK LIGAND  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: palin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,044  
FILING DATE: 05-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/277722  
FILING DATE: 20-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 90203  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3969 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-436-044-5

Alignment Scores:  
Pred. No.: 0.338 Length: 3969  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-436-044-5 (1-3969)  
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
Db 2298 GTCCACCGAGACCTGCGCTCTCGAACATCTTAGTC 2333

RESULT 55  
US-08-222-616-23  
Sequence 23, Application US/08222616  
Patent No. 5635177  
GENERAL INFORMATION:  
APPLICANT: Bennett, Brian D.  
APPLICANT: Goeddel, David  
APPLICANT: Lee, James M.  
APPLICANT: Matthews, William  
APPLICANT: Tsai, Siao Ping  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: palin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222,616  
FILING DATE: 4-APR-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00586  
FILING DATE: 22-JAN-1993  
APPLICATION NUMBER: 07/826935  
FILING DATE: 22-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER:  
REFERENCE/DOCKET NUMBER: 821P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3969 bases  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-222-616-23

Alignment Scores:

Pred. No.:	0.338	Length:	3969
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	1	Gaps:	0

US-09-397-967-16 (1-1099) x US-08-222-616-23 (1-3969)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952  
|||||

Db 2298 GTCCACCGAGACTGCTGCTGCAACATCTAGTC 2333

RESULT 56  
US-08-436-054-5  
Sequence 5, Application US/08436054  
Patent No. 5864020  
GENERAL INFORMATION:  
APPLICANT: Bennett, Brian D.  
APPLICANT: Matthews, William  
TITLE OF INVENTION: HTR LIGAND  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/436.054  
FILING DATE: 05-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/277722  
FILING DATE: 20-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00.000  
REFERENCE/DOCKET NUMBER: 902D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3969 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-436-054-5

Alignment Scores:

Pred. No.:	0.338	Length:	3969
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	2	Gaps:	0

US-09-397-967-16 (1-1099) x US-08-436-054-5 (1-3969)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952

Db 2298 GTCCACCGAGACTGCTGCTGCAACATCTAGTC 2333  
|||||

RESULT 57  
US-08-446-648-23  
Sequence 23, Application US/08446648  
Patent No. 631302  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Bennett, Brian D.  
APPLICANT: Goeddel, David  
APPLICANT: Lee, James M.  
APPLICANT: Matthews, William  
APPLICANT: Tsai, Siao Ping  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446.648  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/222616  
FILING DATE: 04-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0821P3PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3969 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: linear  
US-08-446-648-23

Alignment Scores:

Pred. No.:	0.338	Length:	3969
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.09%	Indels:	0
DB:	4	Gaps:	0

US-09-397-967-16 (1-1099) x US-08-446-648-23 (1-3969)

OY 941 VALHISARGASPLEUALAALARGASNTLEUVAL 952  
|||||

Db 2298 GTCCACCGAGACTGCTGCTGCAACATCTAGTC 2333

RESULT 58  
PCT-US95-04228-23  
Sequence 23, Application PC/TUS9504228  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Bennett, Brian D.  
APPLICANT: Goeddel, David



APPLICANT: Lee, James M.  
APPLICANT: Mathews, William  
APPLICANT: Tsai, Siao Ping  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04228  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/222616  
FILING DATE: 04-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Wendy M. Lee  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 821P3PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3969 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US95-04228-23

Alignment Scores:  
Pred. No.: 0.338 Length: 3969  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0  
DB: 0

US-09-397-967-16 (1-1099) x PCT-US95-04228-23 (1-3969)

QY 941 VALHISARGSPLEUALAALARGSNILEUVAL 952  
DB 2298 GTCCACCGAGACCTGCTCGCAACATCTAGTC 2333

RESULT 59  
PCT-US95-08812-5  
Sequence 5, Application PC/TUS9508812  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
TITLE OF INVENTION: HTRK LIGAND  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/08812  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 902PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3969 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US95-08812-5

Alignment Scores:  
Pred. No.: 0.338 Length: 3969  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0  
DB: 0

US-09-397-967-16 (1-1099) x PCT-US95-08812-5 (1-3969)

QY 941 VALHISARGSPLEUALAALARGSNILEUVAL 952  
DB 2298 GTCCACCGAGACCTGCTCGCAACATCTAGTC 2333

RESULT 60  
US-08-348-143-3  
Sequence 3, Application US/08348143  
Patent No. 5506205  
GENERAL INFORMATION:  
APPLICANT: TAJIMA, HISAO  
APPLICANT: KITAGAWA, KOICHIRO  
APPLICANT: OHNO, HIROYUKI  
APPLICANT: UENO, TOSHIO  
TITLE OF INVENTION: A No. 5506205el Polypeptide of Protein p140 and DNAs  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
STREET: 2100 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20037-3202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/348,143  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 315806/1993  
FILING DATE: 24-NOV-1993  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)293-7060  
TELEFAX: (202)293-7860  
TELEX: 6491103

;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4027 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA to mRNA  
;; ORIGINAL SOURCE:  
;; ORGANISM: rat  
;; TISSUE TYPE: skeletal muscle myoblast  
;; CELL LINE: L6  
US-08-348-143-3

Alignment Scores:  
Pred. No.: 0.343 Length: 4027  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-348-143-3 (1-4027)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
Db 2509 GTGCACCGTGACCTGCTGCCGCGACATCCTTGTG 2544

RESULT 61  
US-08-348-143-4  
; Sequence 4, Application US/08348143  
; Patent No. 5506205  
; GENERAL INFORMATION:  
; APPLICANT: TAJIMA, HISAO  
; APPLICANT: KITAGAWA, KOICHIRO  
; APPLICANT: OHNO, HIROYUKI  
; APPLICANT: UENO, TOSHIO  
; TITLE OF INVENTION: A No. 5506205el Polypeptide of Protein p140 and DNAs  
; TITLE OF INVENTION: encoding it  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SUGHRUE, MIOM, ZINN, MACPEAK & SEAS  
; STREET: 2100 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20037-3202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/348.143  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 315806/1993  
; FILING DATE: 24-NOV-1993  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)293-7060  
; TELEFAX: (202)293-7860  
; TELEX: 6491103  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4027 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; ORIGINAL SOURCE:  
; ORGANISM: rat  
; TISSUE TYPE: skeletal muscle myoblast  
; CELL LINE: L6

;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 262..3243  
;; IDENTIFICATION METHOD: by similarity to some other pattern  
US-08-348-143-4

Alignment Scores:  
Pred. No.: 0.343 Length: 4027  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-348-143-4 (1-4027)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
Db 2509 GTGCACCGTGACCTGCTGCCGCGACATCCTTGTG 2544

RESULT 62  
US-08-571-785-3  
; Sequence 3, Application US/08571785  
; Patent No. 5804411  
; GENERAL INFORMATION:  
; APPLICANT: TAJIMA, HISAO  
; APPLICANT: KITAGAWA, KOICHIRO  
; APPLICANT: OHNO, HIROYUKI  
; APPLICANT: UENO, TOSHIO  
; TITLE OF INVENTION: A No. 5804411el Polypeptide of Protein p140  
; TITLE OF INVENTION: and DNAs encoding it  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SUGHRUE, MIOM, ZINN, MACPEAK & SEAS  
; STREET: 2100 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20037-3202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/571.785  
; FILING DATE: 13-DEC-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/348.143  
; FILING DATE: 23-NOV-1994  
; APPLICATION NUMBER: JP 315806/1993  
; FILING DATE: 24-NOV-1993  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)293-7060  
; TELEFAX: (202)293-7860  
; TELEX: 6491103  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4027 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; ORIGINAL SOURCE:  
; ORGANISM: rat  
; TISSUE TYPE: skeletal muscle myoblast  
; CELL LINE: L6  
US-08-571-785-3

Alignment Scores:  
Pred. No.: 0.343 Length: 4027  
Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-571-785-3 (1-4027)

Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
Db 2509 GTGCACCGTGACCTCGCTGCCGCAACATCCTTGTG 2544

## RESULT 63

US-08-571-785-4

Sequence 4, Application US/08571785  
Patent No. 5804411

## GENERAL INFORMATION:

APPLICANT: TAJIMA, HISAO  
APPLICANT: KITAGAWA, KOICHIRO  
APPLICANT: OHNO, HIROYUKI

APPLICANT: UENO, TOSHIO

TITLE OF INVENTION: A No. 5804411 Polypeptide of Protein p140

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/571,785

FILING DATE: 13-DEC-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/348,143

FILING DATE: 23-NOV-1994

APPLICATION NUMBER: JP 315806/1993

FILING DATE: 24-NOV-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 4027 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: rat

TISSUE TYPE: skeletal muscle myoblast

CELL LINE: L6

FEATURE:

NAME/KEY: CDS

LOCATION: 262..3243

IDENTIFICATION METHOD: by similarity to some other pattern

US-08-571-785-4

Alignment Scores:

Pred. No.: 0.343 Length: 4027

Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.09% Indels: 0

DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-571-785-4 (1-4027)

Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
Db 2509 GTGCACCGTGACCTCGCTGCCGCAACATCCTTGTG 2544

## RESULT 64

US-09-192-435-3

Sequence 3, Application US/09192435  
Patent No. 6303320

## GENERAL INFORMATION:

APPLICANT: TAJIMA, HISAO  
APPLICANT: KITAGAWA, KOICHIRO  
APPLICANT: OHNO, HIROYUKI

APPLICANT: UENO, TOSHIO

TITLE OF INVENTION: A No. 6303320 Polypeptide of Protein p140

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/192,435

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/571,785

FILING DATE: 13-DEC-1995

APPLICATION NUMBER: 08/348,143

FILING DATE: 23-NOV-1994

APPLICATION NUMBER: JP 315806/1993

FILING DATE: 24-NOV-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 4027 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: rat

TISSUE TYPE: skeletal muscle myoblast

CELL LINE: L6

US-09-192-435-3

Alignment Scores:

Pred. No.: 0.343 Length: 4027

Score: 12.00 Matches: 12

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.09% Indels: 0

DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-192-435-3 (1-4027)

Oy 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
Db 2509 GTGCACCGTGACCTCGCTGCCGCAACATCCTTGTG 2544

RESULT 65

```

US-09-192-435-4
; Sequence 4, Application US/09192435
; Patent No. 6303320
; GENERAL INFORMATION:
; APPLICANT: TAJIMA, HISAO
; APPLICANT: KITAGAWA, KOICHIRO
; APPLICANT: OHNO, HIROYUKI
; APPLICANT: UENO, TOSHIO
; TITLE OF INVENTION: A No. 6303320el Polypeptide of Protein p140
; TITLE OF INVENTION: and DMS encoding it
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MIOTON, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/192,435
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/571,785
; FILING DATE: 13-DEC-1995
; APPLICATION NUMBER: 08/348,143
; FILING DATE: 23-NOV-1994
; APPLICATION NUMBER: JP 315806/1993
; FILING DATE: 24-NOV-1993
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4027 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: rat
; TISSUE TYPE: skeletal muscle myoblast
; CELL LINE: L6
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 262..3243
; IDENTIFICATION METHOD: by similarity to some other pattern
US-09-192-435-4

Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB

0.343
12.00
100.00%
100.00%
1.09%
4
1
0

Length:
Matches:
Conservative:
Mismatch:
Indels:
Gaps:
4027
12
0
0
0
0

US-09-397-967-16 (1-1099) x US-09-192-435-4 (1-4027)
QY 941 VALHISARGASPLEUVALAALAARGASNILEUVAL 952
| | | | | | | | | | | | | | | | | | | | |
DB 2509 GTGCACCGTGACTTCGCTGCCCGACATCCTTGTG 2544

RESULT 66
US-09-558-340-3
; Sequence 3, Application US/09558340
; Patent No. 6432913

```

```

GENERAL INFORMATION:
APPLICANT:  TAJIMA, HISAO
APPLICANT:  KITAGAWA, KOICHIRO
APPLICANT:  OHNO, HIROYUKI
APPLICANT:  UENO, TOSHIO
TITLE OF INVENTION:  A No. 6432913el Polypeptide of Protein p140
TITLE OF INVENTION:  and dnas encoding it
NUMBER OF SEQUENCES:  16
CORRESPONDENCE ADDRESSES:
ADDRESSEE:  SUGHRUE, MIION, ZINN, MACPEAK & SEAS
STREET:  2100 Pennsylvania Avenue, N.W.
CITY:  Washington
STATE:  D.C.
COUNTRY:  U.S.A.
ZIP:  20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE:  Floppy disk
COMPUTER:  IBM PC compatible
OPERATING SYSTEM:  PC-DOS/MS-DOS
SOFTWARE:  Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER:  US/09/558,340
FILING DATE:  26-APR-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:  09/192,435
FILING DATE:  08-JAN-1998
APPLICATION NUMBER:  08/571,785
FILING DATE:  13-DEC-1995
APPLICATION NUMBER:  08/348,143
FILING DATE:  23-NOV-1994
APPLICATION NUMBER:  JP 315806/1993
FILING DATE:  24-NOV-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE:  (202)293-7060
TELEFAX:  (202)293-7860
TELEX:  6491103
INFORMATION FOR SEQ ID NO:  3:
SEQUENCE CHARACTERISTICS:
LENGTH:  4027 base pairs
TYPE:  nucleic acid
STRANDEDNESS:  single
TOPOLOGY:  linear
MOLECULE TYPE:  cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM:  rat
TISSUE TYPE:  skeletal muscle myoblast
CELL LINE:  L6
US-09-558-340-3

Alignment Scores:
Pred. No.:          0.343          Length:      4027
Score:              12.00          Matches:     12
Percent Similarity: 100.00%       Conservative: 0
Best Local Similarity: 100.00%   Mismatches:  0
Query Match:        1.09%         Indels:      0
                                   Gaps:        0
DB:
US-09-397-967-16 (1-1099) x US-09-558-340-3 (1-4027)
Oy  941 ValHisArgAspLeuAlaIleArgAsnIleLeuVal 952
    |||||||
Db  2509 GTGCACCGTACCTCGCTGCCCGCACACATCCTGTGC 2544
    |||||||

RESULT 67
US-09-558-340-4
; Sequence 4, Application US/09558340
; Patent No. 6432913
;
; GENERAL INFORMATION:
; APPLICANT:  TAJIMA, HISAO
; APPLICANT:  KITAGAWA, KOICHIRO
; APPLICANT:  OHNO, HIROYUKI
; APPLICANT:  UENO, TOSHIO

```

;; TITLE OF INVENTION: A No. 6432913el Polypeptide of Protein p140  
;; TITLE OF INVENTION: and DNAs encoding it  
;; NUMBER OF SEQUENCES: 16  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: SUGHRUE, MIOW, ZINN, MACPHEAK & SEAS  
;; STREET: 2100 Pennsylvania Avenue, N.W.  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: U.S.A.  
;; ZIP: 20037-3202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/558,340  
;; FILING DATE: 26-APR-2000  
;; CLASSIFICATION:  
;; PRIORITY APPLICATION DATA:  
;; APPLICATION NUMBER: 09/192,435  
;; FILING DATE: 08-JAN-1998  
;; APPLICATION NUMBER: 08/571,785  
;; FILING DATE: 13-DEC-1995  
;; APPLICATION NUMBER: 08/348,143  
;; FILING DATE: 23-NOV-1994  
;; APPLICATION NUMBER: JP 315806/1993  
;; FILING DATE: 24-NOV-1993  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202)293-7060  
;; TELEFAX: (202)293-7860  
;; TELEX: 6491103  
;; INFORMATION FOR SEQ. ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4027 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA to mRNA  
;; ORIGINAL SOURCE:  
;; ORGANISM: rat  
;; TISSUE TYPE: skeletal muscle myoblast  
;; CELL LINE: L6  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 262..3243  
;; IDENTIFICATION METHOD: by similarity to some other pattern  
US-09-558-340-4  
  
Alignment Scores:  
Pred. No.: 0.343 Length: 4027  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 4 Gaps: 0  
  
US-09-397-967-16 (1-1099) x US-09-558-340-4 (1-4027)  
QY 941 VALHISARGASPLEUALAIAARGASNIIELEVAL 952  
DB 2509 GTGCACCGTGACCTGCTGCCGACACATCTTGTC 2344  
  
RESULT 68  
US-08-162-809-17  
; Sequence 17, Application US/08162809  
; Patent No. 5457048  
; GENERAL INFORMATION:  
; APPLICANT: Pasquale, Elena B.  
; APPLICANT: Sajjadi, Fereydoun G.  
; TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE  
; NUMBER OF SEQUENCES: 26

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: CAMPBELL AND FLORES  
;; STREET: 4370 La Jolla Village Drive, Suite 700  
;; CITY: San Diego  
;; STATE: California  
;; COUNTRY: United States of America  
;; ZIP: 92122  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/162,809  
;; FILING DATE:  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Campbell, Cathryn A.  
;; REGISTRATION NUMBER: P-LJ 9503  
;; REFERENCE/DOCKET NUMBER: 31,815  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 535-9001  
;; TELEFAX: (619) 535-8949  
;; INFORMATION FOR SEQ. ID NO: 17:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4049 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: both  
;; TOPOLOGY: linear  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 10..2994  
US-08-162-809-17  
  
Alignment Scores:  
Pred. No.: 0.344 Length: 4049  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 1 Gaps: 0  
  
US-09-397-967-16 (1-1099) x US-08-162-809-17 (1-4049)  
QY 941 VALHISARGASPLEUALAIAARGASNIIELEVAL 952  
DB 2263 GTGCACCGGACCTGCTGCCGACACATCTTGTC 2298  
  
RESULT 69  
US-08-162-809-11  
; Sequence 11, Application US/08162809  
; Patent No. 5457048  
; GENERAL INFORMATION:  
; APPLICANT: Pasquale, Elena B.  
; APPLICANT: Sajjadi, Fereydoun G.  
; TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CAMPBELL AND FLORES  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/162,809  
; FILING DATE:

CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 9503  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4097 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 10..3042  
US-08-162-809-11  
Alignment Scores:  
Pred. No.: 0.348 Length: 4097  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0  
DB: 1  
US-09-397-967-16 (1-1099) x US-08-162-809-11 (1-4097)  
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
Db 2311 GTGCACCGGAGACTGTGCTCCGCAACATCTGTGTC 2346  
RESULT 70  
US-08-368-776A-1  
Sequence 1, Application US/08368776A  
Patent No. 6300482  
GENERAL INFORMATION:  
APPLICANT: Ciossek, Thomas  
APPLICANT: Ullrich, Axel  
TITLE OF INVENTION: METHODS FOR DIAGNOSIS  
TITLE OF INVENTION: AND TREATMENT OF MDK1  
TITLE OF INVENTION: SIGNAL TRANSDUCTION  
TITLE OF INVENTION: DISORDERS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/368,776A  
FILING DATE: January 3, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/007  
none

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4304 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic  
US-08-368-776A-1  
Alignment Scores:  
Pred. No.: 0.364 Length: 4304  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
Gaps: 0  
DB: 4  
US-09-397-967-16 (1-1099) x US-08-368-776A-1 (1-4304)  
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952  
Db 2495 GTTCACAGGAGACTTGCAGCGCGCAACATCTGTGTC 2530  
PCT-US96-00419-1  
PCT-US96-00419-1  
Sequence 1, Application PC/TUS9600419  
GENERAL INFORMATION:  
APPLICANT: Thomas Ciossek, Axel Ullrich, Birgit  
APPLICANT: Millaueer  
TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF  
TITLE OF INVENTION: MDK1 SIGNAL TRANSDUCTION DISORDERS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/00419  
FILING DATE: January 3, 1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4304  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic  
PCT-US96-00419-1  
none

```
Alignment Scores:
Pred. No.: 0.364 Length: 4304
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US96-00419-1 (1-4304)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2495 GTTCACAGGACCTTGCGAGCGCGCAACATCTTGTC 2530

RESULT 72
PCT-US93-06251-34
Sequence 34, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wickstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
Stereospecific Alkylphosphonates and Arylphosphonates
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DIGILIO, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 4508 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-34

Alignment Scores:
Pred. No.: 0.38 Length: 4508
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US93-06251-34 (1-4508)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2542 GTTCATCGGACCTTGCGAGCGCGCAACATCTTGTA 2577

RESULT 73
US-08-449-645A-16
Sequence 16, Application US/08449645A

Patent No. 5981245
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,645A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 4529 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 186..3182
US-08-449-645A-16

Alignment Scores:
Pred. No.: 0.381 Length: 4529
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-449-645A-16 (1-4529)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
DB 2448 GTTCACAGGACCTTGCGAGCTGCGCAATATCTTGTC 2483

RESULT 74
US-08-702-367A-16
Sequence 16, Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
```

```

APPLICATION NUMBER: US/08/702,367A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 4529 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 186..3182
US-08-702-367A-16

Alignment Scores:
Pred. No.: 0.381 Length: 4529
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB:

US-09-397-967-16 (1-1099) x US-08-702-367A-16 (1-4529)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2448 GTTCACAGGAGCTTCGACCTCGCATATTTCTGTGC 2483

RESULT 75
PCT-US95-04681-16
Sequence 16, Application PC/TUS9504681
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: Eph-like Receptor Protein Tyrosine
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehaviiland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04681
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 4529 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 186..3182
PCT-US95-04681-16

Alignment Scores:
Pred. No.: 0.381 Length: 4529

```

```

Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB:

US-09-397-967-16 (1-1099) x PCT-US95-04681-16 (1-4529)
OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952
Db 2448 GTTCACAGGAGCTTCGACCTCGCATATTTCTGTGC 2483

RESULT 76
US-08-857-076-11
Sequence 11, Application US/08857076C
GENERAL INFORMATION:
APPLICANT: Ruvkun, Gary
APPLICANT: Kimura, Koutarou
APPLICANT: Patterson, Garth
APPLICANT: Ogil, Scott
APPLICANT: Paradis, Suzanne
APPLICANT: Tissenbaum, Heidi
APPLICANT: Morris, Jason
APPLICANT: Kowsek, Allison
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
FILE REFERENCE: 00786/351001
CURRENT FILING DATE: 1997-05-15
NUMBER OF SEQ ID NOS: 114
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 5816
TYPE: DNA
ORGANISM: Caenorhabditis elegans
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(5816)
OTHER INFORMATION: n = A,T,C or G
US-08-857-076-11

Alignment Scores:
Pred. No.: 0.479 Length: 5816
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.09% Indels: 0
Gaps: 0
DB:

US-09-397-967-16 (1-1099) x US-08-857-076-11 (1-5816)
OY 1003 SerAspValTrpSerPheGlyValIleuVal 1014
Db 4455 TGTGATGTTGGAGCTTCGAGTGTCTCTATGA 4490

RESULT 77
US-09-741-154-3
Sequence 3, Application US/09741154
GENERAL INFORMATION:
APPLICANT: BEASLEY, Ellen M. et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: C1001061
CURRENT APPLICATION NUMBER: US/09/741,154
CURRENT FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 16389
TYPE: DNA

```



ORGANISM: Human  
US-09-741-154-3

Alignment Scores:  
Pred. No.: 1.23 Length: 16389  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-741-154-3 (1-16389)

RESULT 78  
US-08-222-616-7  
Sequence 7, Application US/08222616  
Patent No. 5635177  
GENERAL INFORMATION:  
APPLICANT: Bennett, Brian D.  
APPLICANT: Goeddel, David  
APPLICANT: Lee, James M.  
APPLICANT: Matthews, William  
APPLICANT: Tsai, Siao Ping  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: paln (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222,616  
FILING DATE: 4-APR-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00586  
FILING DATE: 22-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/826935  
FILING DATE: 22-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER:  
REFERENCE/DOCKET NUMBER: 821P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 147 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-222-616-7

Alignment Scores:  
Pred. No.: 0.137 Length: 147  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-222-616-7 (1-147)

RESULT 79  
US-08-876-882-3  
Sequence 3, Application US/08876882  
Patent No. 5981201  
GENERAL INFORMATION:  
APPLICANT: Avraham, Hava  
APPLICANT: Groopman, Jerome E.  
TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds P.C.  
STREET: Two Mallitia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173-4799

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/876,882  
FILING DATE: 16-JUN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/035,228  
FILING DATE: 08-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Doreen, Hoyle M.  
REGISTRATION NUMBER: 36,361  
REFERENCE/DOCKET NUMBER: NEDH97-01PA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 781-861-6240  
TELEFAX: 781-861-9540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 147 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-876-882-3

Alignment Scores:  
Pred. No.: 0.137 Length: 147  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-876-882-3 (1-147)

RESULT 80  
US-08-446-648-7  
Sequence 7, Application US/08446648  
Patent No. 6331302

GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Bennett, Brian D.  
APPLICANT: Goeddel, David  
APPLICANT: Lee, James M.  
APPLICANT: Matthews, William  
APPLICANT: Tsai, Siao Ping  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,648  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/222616  
FILING DATE: 04-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0821P3PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 147 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
US-08-446-648-7

Alignment Scores:  
Pred. No.: 0.137 Length: 147  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-446-648-7 (1-147)

QY 942 HisArgAspLeuAlaAlaArgAsnIleuVal 952  
|||||  
DB 10 CACAGAGACTAGCAGCAGCAGCATCTGCTC 42

RESULT 81  
US-09-315-928-3  
Sequence 3, Application US/09315928  
Patent No. 6368796  
GENERAL INFORMATION:  
APPLICANT: Avraham, Hava  
APPLICANT: Groopman, Jerome E.  
TITLE OF INVENTION: METHODS OF DETECTION AND TREATMENT OF  
FILE REFERENCE: NEDH97-01P4Z  
CURRENT APPLICATION NUMBER: US/09/315,928  
CURRENT FILING DATE: 1999-05-20  
PRIOR APPLICATION NUMBER: US 08/876,882  
PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: US 60/035,228  
PRIOR FILING DATE: 1997-01-08  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 147  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(150)  
US-09-315-928-3

Alignment Scores:  
Pred. No.: 0.137 Length: 147  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967-16 (1-1099) x US-09-315-928-3 (1-147)

QY 942 HisArgAspLeuAlaAlaArgAsnIleuVal 952  
|||||  
DB 10 CACAGAGACTAGCAGCAGCAGCATCTGCTC 42

RESULT 82  
PCT-US95-04228-7  
Sequence 7, Application PC/TUS9504228  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Bennett, Brian D.  
APPLICANT: Goeddel, David  
APPLICANT: Lee, James M.  
APPLICANT: Matthews, William  
APPLICANT: Tsai, Siao Ping  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 KB floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04228  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/222616  
FILING DATE: 04-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Wendy M. Lee  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 821P3PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 147 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

PCT-US95-04228-7

## Alignment Scores:

Pred. No.:	0.137	Length:	147
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.00%	Indels:	0
DB:	5	Gaps:	0

US-09-397-967-16 (1-1099) x PCT-US95-04228-7 (1-147)

QY 942 HisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 10 CACAGAGACCTAGCAGCAGCAGCATCTGTC 42

## RESULT 83

US-08-278-089A-3

Sequence 3, Application US/08278089A

Patent No. 5681714

## GENERAL INFORMATION:

APPLICANT: Breilman, Martin L.

APPLICANT: Rosant, Janet

APPLICANT: Dumont, Daniel J.

APPLICANT: Yamaguchi, Terry P.

TITLE OF INVENTION: No. 5681714el Receptor Tyrosine Kinase

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bereskin &amp; Parr

STREET: 40 King Street West

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/278,089A

FILING DATE: 20-JUL-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Kurdyak, Linda M.

REGISTRATION NUMBER: 34,971

REFERENCE/DOCKET NUMBER: 3153-111

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 364-7311

TELEFAX: (416) 361-1398

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1590 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Mus musculus

DEVELOPMENTAL STAGE: Embryo

IMMEDIATE SOURCE:

LIBRARY: murine embryonic lambda gt10 cDNA library

CLONE: 1.6kb clone

POSITION IN GENOME:

CHROMOSOME/SEGMENT: 4

MAP POSITION: Between the brown and pmv-23 loci

FEATURE:

NAME/KEY: CDS

LOCATION: 1..903

US-08-278-089A-3

## Alignment Scores:

Pred. No.:	1.19	Length:	1590
------------	------	---------	------

Score: 11.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.00%

DB: 1

US-09-397-967-16 (1-1099) x US-08-278-089A-3 (1-1590)

QY 942 HisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 415 CACAGGACCTGGCTGCCAGAACATTTTACTT 447

## RESULT 84

US-07-934-393B-3

Sequence 3, Application US/07934393B

Patent No. 546596

## GENERAL INFORMATION:

APPLICANT: BREITMAN, MARTIN L.

APPLICANT: DUMONT, DANIEL

APPLICANT: GRADWOHL, GERARD G.

TITLE OF INVENTION: TISSUE SPECIFIC TRANSCRIPTIONAL

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: BERESKIN &amp; PARR

STREET: 40 King Street West

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/934,393B

FILING DATE: 25-AUG-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kurdyak, Linda M.

REGISTRATION NUMBER: 34,971

REFERENCE/DOCKET NUMBER: 3153-64

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 354-7311

TELEFAX: (416) 361-1398

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1601 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Mus pahari

DEVELOPMENTAL STAGE: Embryo

IMMEDIATE SOURCE:

LIBRARY: murine embryonic lambda gt10 cDNA library

CLONE: 1.6kb clone

POSITION IN GENOME:

CHROMOSOME/SEGMENT: 4

MAP POSITION: Between the brown and pmv-23 loci

FEATURE:

NAME/KEY: CDS

LOCATION: 1..903

US-07-934-393B-3

## Alignment Scores:

Pred. No.:	1.2	Length:	1601
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.00%	Indels:	0

DB:	1	Caps:
US-09-397-967-16 (1-1099) x US-07-934-392B-3 (1-1601)		
Qy	942 HisArgAspLeuAlaIalargsnlleleual	952
Dd	415 CACGGAGCTGCTGCCAGAACATTTCATTTACTT	447

```

US-08-838-957A-3
; Sequence 3, Application US/08838957A
; Patent No. 5998187
; GENERAL INFORMATION:
; APPLICANT: Breitman, Martin L.
; APPLICANT: Rossant, Janet
; APPLICANT: Dumont, Daniel J.
; APPLICANT: Yamauchi, Terry P.
; TITLE OF INVENTION: No. 5998187e1 Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3T2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,957A
; FILING DATE: 23-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurdzyk, Linda M.
; REGISTRATION NUMBER: 34,971
; REFERENCE/DOCKET NUMBER: 3153-212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Mus musculus
; DEVELOPMENTAL STAGE: Embryo
; IMMEDIATE SOURCE:
; LIBRARY: Murine embryonic lambda gt10 cDNA library
; CLONE: 1.6kb clone
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 4
; MAP POSITION: Between the brown and pmv-23 loci
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..903
; US-08-838-957A-3

Alignment Scores:
Prod. No.: 1.2 Length: 1601
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.00% Indels: 0
                2 Gaps: 0
US-09-397-967-16 (1-1099) x US-08-838-957A-3 (1-1601)

```

```
Oy      942 HisARGSPLeuaLaalaARgasmIleueVal    952  
        |||||  
Db      415 CACAGGGACTGCGTGCCAGAACAATTATGT     447
```

```

US-08-162-809-21
; Sequence 21, Application US/08162809
; Patent No. 5457048
; GENERAL INFORMATION:
; APPLICANT: Pasquale, Elena B.
; TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES, AND METHODS OF USE
; NUCLEOTIDE SEQUENCES, AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/162, 809
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELEPHONE: (619) 535-8941
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 3056 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 2..2131
US-08-162-809-21
Alignment Scores:
Pred. No.: 2,16 Length: 3056
Score: 11,00 Matches: 11
Percent Similarity: 100,00% Conservative: 0
Best Local Similarity: 100,00% Mismatches: 0
Query Match: 1,00% Indels: 0
Gaps: 0
DB: 1
US-09-397-967-16 (1-1099) x US-08-162-809-21 (1-3056)
Oy 941 ValHisArgAspLeuAlaIalaArgAsnIleLeu 951
| | | | | | | | | | | | | | | | | | | | | | | |
Db 1445 GTACACAGACCTCGCTGCAGGAATATCTTC 1477
RESULT 87
; Sequence 3, Application US/08162809
; Patent No. 5457048
; GENERAL INFORMATION:
; APPLICANT: Pasquale, Elena B.
; TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES, AND METHODS OF USE
; NUCLEOTIDE SEQUENCES, AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:

```

ADDRESSEE: CAMPBELL AND FLORES  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States of America  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/162,809  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 9503  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3059 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..2167  
US-08-162-809-3

Alignment Scores:  
Pred. No.: 2.16 Length: 3059  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-3 (1-3059)

QY 941 VALHISARGASPLEUALAALARGASNIIELEU 951  
|||||  
DB 1445 GTACACAGACCTCGCTGCCAGAAATATCTTC 1477

RESULT 88  
US-08-162-809-19  
Sequence 19, Application US/08162809  
Patent No. 5457048  
GENERAL INFORMATION:  
APPLICANT: Pasquale, Elena B.  
APPLICANT: Sajjad, Fereydoon G.  
TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CAMPBELL AND FLORES  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States of America  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/162,809  
FILING DATE:  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 9503  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3125 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..2233  
US-08-162-809-19

Alignment Scores:  
Pred. No.: 2.2 Length: 3125  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-19 (1-3125)

QY 941 VALHISARGASPLEUALAALARGASNIIELEU 951  
|||||  
DB 1511 GTACACAGACCTCGCTGCCAGAAATATCTTC 1543

RESULT 89  
US-08-167-919A-9  
Sequence 9, Application US/08167919A  
Patent No. 5674691  
GENERAL INFORMATION:  
APPLICANT: Boyd, Andrew W.  
APPLICANT: Simpson, Richard J.  
APPLICANT: Wicks, Ian  
APPLICANT: Ward, Larry D.  
APPLICANT: Wilkinson, David  
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE  
TITLE OF INVENTION: AND USE THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/167,919A  
FILING DATE: 18-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PK6841 (AU)  
FILING DATE: 21-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PK9992 (AU)  
FILING DATE: 12-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU92/00294  
FILING DATE: 19-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 9159  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3132 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 100..3048  
US-08-167-919A-9

Alignment Scores:  
Pred. No.: 2.21 Length: 3132  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
Gaps: 0  
DB:

US-09-397-967-16 (1-1099) x US-08-167-919A-9 (1-3132)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeu 951  
|||||  
DB 2326 GTTCACCGAGACCTGCTGCTCGGAACATCTTG 2358

RESULT 90  
US-08-715-106-9  
Sequence 9, Application US/08715106  
Patent No. 6020306  
GENERAL INFORMATION:  
APPLICANT: Boyd, Andrew W.  
APPLICANT: Simpson, Richard J.  
APPLICANT: Wicks, Ian  
APPLICANT: Ward, Larry D.  
APPLICANT: Wilkinson, David  
TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE TYROSINE KINASE  
TITLE OF INVENTION: AND USE THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/715,106  
FILING DATE: 18-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/167,919  
FILING DATE: 18-APR-1994  
APPLICATION NUMBER: PK6841 (AU)  
FILING DATE: 21-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PK9992 (AU)  
FILING DATE: 12-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU92/00294  
FILING DATE: 19-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9159  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3132 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 100..3048  
US-08-715-106-9

Alignment Scores:  
Pred. No.: 2.21 Length: 3132  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
Gaps: 0  
DB:

US-09-397-967-16 (1-1099) x US-08-715-106-9 (1-3132)

OY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeu 951  
|||||  
DB 2326 GTTCACCGAGACCTGCTGCTCGGAACATCTTG 2358

RESULT 91  
US-08-449-645A-12  
Sequence 12, Application US/08449645A  
Patent No. 5981245  
GENERAL INFORMATION:  
APPLICANT: Fox, Gary M.  
TITLE OF INVENTION: EPH-like Receptor Protein Tyrosine  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Patent Operations/RBW  
STREET: 1840 Dehavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,645A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Winter, Robert B.  
REFERENCE/DOCKET NUMBER: A-287  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3162 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2976  
US-08-449-645A-12  
Alignment Scores:

Pred. No.: 2.23 Length: 3162  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-449-645A-12 (1-3162)

QY 941 VALHISARGASPLEUAlaAlaArgAsnIleLeu 951

Db 2251 GTGCATAGAGATCTTGCTGCCAGAAACATCTTA 2283

RESULT 92

US-08-702-367A-12

Sequence 12, Application US/08702367A

Patent No. 5981246

GENERAL INFORMATION:

APPLICANT: FOX, Gary M.

TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Patent Operations/RBM

STREET: 1840 Dehavilland Drive

CITY: Thousand Oaks

STATE: California

COUNTRY: USA

ZIP: 91320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/702.367A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Winter, Robert B.

REFERENCE/DOCKET NUMBER: A-287

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 3162 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2976

US-08-702-367A-12

Alignment Scores:

Pred. No.: 2.23

Score: 11.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.00%

DB: 2

US-09-397-967-16 (1-1099) x US-08-702-367A-12 (1-3162)

QY 941 VALHISARGASPLEUAlaAlaArgAsnIleLeu 951

Db 2251 GTGCATAGAGATCTTGCTGCCAGAAACATCTTA 2283

RESULT 93

PCT-US95-04681-12

Sequence 12, Application PC/TUS9504681

GENERAL INFORMATION:

APPLICANT: FOX, Gary M.

TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine

TITLE OF INVENTION: Kinases  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Patent Operations/RBM  
STREET: 1840 Dehavilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA

ZIP: 91320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/04681.

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Winter, Robert B.

REFERENCE/DOCKET NUMBER: A-287

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 3162 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2976

PCT-US95-04681-12

Alignment Scores:

Pred. No.: 2.23

Score: 11.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.00%

DB: 5

US-09-397-967-16 (1-1099) x PCT-US95-04681-12 (1-3162)

QY 941 VALHISARGASPLEUAlaAlaArgAsnIleLeu 951

Db 2251 GTGCATAGAGATCTTGCTGCCAGAAACATCTTA 2283

RESULT 94

US-08-162-809-15

Sequence 15, Application US/08162809

Patent No. 5457048

GENERAL INFORMATION:

APPLICANT: Pasquale, Elena B.

TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: CAMPBELL AND FLORES

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: United States of America

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/162,809

FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 9503  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9901  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3254 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 32..2980  
US-08-162-809-15

Alignment Scores:  
Pred. No.: 2.29 Length: 3254  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-162-809-15 (1-3254)  
QY 941 VALHISARGASPLEUALAALARGASNILEU 951  
DB 2258 GTCCACGAGATCTGCTGCTGTAATATACCTC 2290

RESULT 95  
US-08-469-537A-102  
Sequence 102, Application US/08469537A  
Patent No. 5843749  
GENERAL INFORMATION:  
APPLICANT: Maisonsplere, et al.  
TITLE OF INVENTION: EHK AND FOR TYROSINE  
TITLE OF INVENTION: KINASES  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESS: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: NY  
COUNTRY: U.S.A.  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,537A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/406,247  
FILING DATE: 17-MAR-1995  
APPLICATION NUMBER: USSN 08/144,992  
FILING DATE: 28-OCT-1993  
APPLICATION NUMBER: USSN 07/736,559  
FILING DATE: 26-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempster, Ph.D., Gall M  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: REG 070C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
TELEX:  
INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:  
LENGTH: 3906 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 476..3493  
US-08-469-537A-102

Alignment Scores:  
Pred. No.: 2.7 Length: 3906  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
DB: 2 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-469-537A-102 (1-3906)  
QY 941 VALHISARGASPLEUALAALARGASNILEU 951  
DB 2870 GTCCACGAGAGACCTTGCTGCTAGAACATCTTA 2902

RESULT 96  
US-08-323-474-1  
Sequence 1, Application US/08323474  
Patent No. 5447860  
GENERAL INFORMATION:  
APPLICANT: Ziegler, Steven F.  
TITLE OF INVENTION: NOVEL TYROSINE KINASE  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/323,474  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/905,600  
FILING DATE: 26-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Seese, Cathryn A.  
REGISTRATION NUMBER: 32,172  
REFERENCE/DOCKET NUMBER: 2609  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206) 233-0644  
TELEX: 756822  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4138 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 149..3523  
US-08-323-474-1

Alignment Scores:



Pred. No.: 1 2.84 Length: 4138  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.00% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-323-474-1 (1-4138)

OY 942 HlsarqaspleuAlaAlaArqasnlleuVal 952

Db 3032 CACAGGATCTGCTGCCAGAAACATTTAGTT 3064

RESULT 97

PCT-US93-06093-1

Sequence 1, Application PC/TUS9306093

GENERAL INFORMATION:

APPLICANT: Ziegler, Steven F.

TITLE OF INVENTION: NOVEL TYROSINE KINASE

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/06093

FILING DATE: 19930625

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/905,600

FILING DATE: 26-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Seese, Kathryn A.

REGISTRATION NUMBER: 32,172

REFERENCE/DOCKET NUMBER: 2609

TELEPHONE: (206) 587-0430

TELEFAX: (206) 233-0644

TELEX: 756822

INFORMATION FOR SEQ. ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4138 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE:

NAME/KEY: CDS

LOCATION: 149..3523

PCT-US93-06093-1

Alignment Scores:

Pred. No.: 2.84 Length: 4138

Score: 11.00 Matches: 11

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.00% Indels: 0

DB: 5 Gaps: 0

US-09-397-967-16 (1-1099) x PCT-US93-06093-1 (1-4138)

OY 942 HlsarqaspleuAlaAlaArqasnlleuVal 952

Db 3032 CACAGGATCTGCTGCCAGAAACATTTAGTT 3064

RESULT 98

US-08-442-248-1

Sequence 1, Application US/08442248

Patent No. 5759863

GENERAL INFORMATION:

APPLICANT: Caras, Ingrid W.

TITLE OF INVENTION: AL-1 Neurotrophic Factor

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patlo (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/442,248

FILING DATE: 15-MAY-1995

CLASSIFICATION: A35

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/330128

FILING DATE: 27-OCT-1994

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, Timothy E.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: 920C4

TELEPHONE: 415/225-8674

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ. ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4165 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-442-248-1

Alignment Scores:

Pred. No.: 2.86 Length: 4165

Score: 11.00 Matches: 11

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.00% Indels: 0

DB: 1 Gaps: 0

US-09-397-967-16 (1-1099) x US-08-442-248-1 (1-4165)

OY 941 ValHlsarqaspleuAlaAlaArqasnlleuVal 951

Db 2602 GTGCACAGACCTGCTGCTAGAAACATCTTA 2634

RESULT 99

US-08-440-815-1

Sequence 1, Application US/08440815

Patent No. 5798448

GENERAL INFORMATION:

APPLICANT: Caras, Ingrid W.

TITLE OF INVENTION: AL-1 Neurotrophic Factor

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genetech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440, 815  
FILING DATE: 15-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/330128  
FILING DATE: 27-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Toricha, Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: 920C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8674  
TELEFAX: 415/952-9881  
TELEX: 910/571-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4165 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-440-815-1

Alignment Scores:	
Pred. No.:	2.86
Score:	11.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	1.00%
DB:	1
Length:	4165
Matches:	11
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

```

US-09-397-967-16. (1-1099) x US-08-440-815-1 (1-14165)
Qy 941 VALHSATgAspleaialaargAsnllleu 951
    |||||
    |||||
    |||||
Db 2602 GNGCACAAGACCTTGGCTGCTAGAAACATCTTA 2634

```

RESULT 100  
 US-08-486-449-1  
 Sequence 1, Application US/08486449  
 Patent No. 6280732  
 GENERAL INFORMATION:  
 APPLICANT: Caras, Ingrid W.  
 APPLICANT: Wlaslow, John W.  
 TITLE OF INVENTION: AL-1 Neurotrophic Factor  
 NUMBER OF SEQUENCES: 18  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genentech, Inc.  
 STREET: 460 Point San Bruno Blvd  
 CITY: South San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94080  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: patlin (Genentech)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/486,449  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/330128  
 FILING DATE: 27-OCT-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Torchia, Timothy E.

```

      1  REGISTRATION NUMBER: 36,700
      2  REFERENCE/DOCKET NUMBER: P0920P1
      3  TELECOMMUNICATION INFORMATION:
      4  TELEPHONE: 415/225-8674
      5  TELEFAX: 415/952-9881
      6  TELEX: 910/371-7168
      7  INFORMATION FOR SEQ ID NO: 1:
      8  SEQUENCE CHARACTERISTICS:
      9  LENGTH: 4165 bases
     10  TYPE: nucleic acid
     11  STRANDEDNESS: single
     12  TOPOLOGY: linear
     13  US-08-486-449-1
     14
     15  Alignment Scores:
     16  Pred. No.:
     17  Score: 2.86
     18  Percent Similarity: 11.00
     19  Best Local Similarity: 100.00%
     20  Query Match: 1.00%
     21  DB: 4
     22
     23  US-08-397-967-16 (1-1099) x US-08-486-449-1 (1-4165)
     24
     25  QY 941 ValHisArgSpLeuAlaAlaArgAsnIleLeu 951
     26  |||||||
     27  |||||||
     28  |||||||
     29  |||||||
     30  |||||||
     31  |||||||
     32  |||||||
     33  |||||||
     34  |||||||
     35  |||||||
     36  |||||||
     37  |||||||
     38  |||||||
     39  |||||||
     40  |||||||
     41  |||||||
     42  |||||||
     43  |||||||
     44  |||||||
     45  |||||||
     46  |||||||
     47  |||||||
     48  |||||||
     49  |||||||
     50  |||||||
     51  |||||||
     52  |||||||
     53  |||||||
     54  |||||||
     55  |||||||
     56  |||||||
     57  |||||||
     58  |||||||
     59  |||||||
     60  |||||||
     61  |||||||
     62  |||||||
     63  |||||||
     64  |||||||
     65  |||||||
     66  |||||||
     67  |||||||
     68  |||||||
     69  |||||||
     70  |||||||
     71  |||||||
     72  |||||||
     73  |||||||
     74  |||||||
     75  |||||||
     76  |||||||
     77  |||||||
     78  |||||||
     79  |||||||
     80  |||||||
     81  |||||||
     82  |||||||
     83  |||||||
     84  |||||||
     85  |||||||
     86  |||||||
     87  |||||||
     88  |||||||
     89  |||||||
     90  |||||||
     91  |||||||
     92  |||||||
     93  |||||||
     94  |||||||
     95  |||||||
     96  |||||||
     97  |||||||
     98  |||||||
     99  |||||||
    100  |||||||
    101  |||||||
    102  |||||||
    103  |||||||
    104  |||||||
    105  |||||||
    106  |||||||
    107  |||||||
    108  |||||||
    109  |||||||
    110  |||||||
    111  |||||||
    112  |||||||
    113  |||||||
    114  |||||||
    115  |||||||
    116  |||||||
    117  |||||||
    118  |||||||
    119  |||||||
    120  |||||||
    121  |||||||
    122  |||||||
    123  |||||||
    124  |||||||
    125  |||||||
    126  |||||||
    127  |||||||
    128  |||||||
    129  |||||||
    130  |||||||
    131  |||||||
    132  |||||||
    133  |||||||
    134  |||||||
    135  |||||||
    136  |||||||
    137  |||||||
    138  |||||||
    139  |||||||
    140  |||||||
    141  |||||||
    142  |||||||
    143  |||||||
    144  |||||||
    145  |||||||
    146  |||||||
    147  |||||||
    148  |||||||
    149  |||||||
    150  |||||||
    151  |||||||
    152  |||||||
    153  |||||||
    154  |||||||
    155  |||||||
    156  |||||||
    157  |||||||
    158  |||||||
    159  |||||||
    160  |||||||
    161  |||||||
    162  |||||||
    163  |||||||
    164  |||||||
    165  |||||||
    166  |||||||
    167  |||||||
    168  |||||||
    169  |||||||
    170  |||||||
    171  |||||||
    172  |||||||
    173  |||||||
    174  |||||||
    175  |||||||
    176  |||||||
    177  |||||||
    178  |||||||
    179  |||||||
    180  |||||||
    181  |||||||
    182  |||||||
    183  |||||||
    184  |||||||
    185  |||||||
    186  |||||||
    187  |||||||
    188  |||||||
    189  |||||||
    190  |||||||
    191  |||||||
    192  |||||||
    193  |||||||
    194  |||||||
    195  |||||||
    196  |||||||
    197  |||||||
    198  |||||||
    199  |||||||
    200  |||||||
    201  |||||||
    202  |||||||
    203  |||||||
    204  |||||||
    205  |||||||
    206  |||||||
    207  |||||||
    208  |||||||
    209  |||||||
    210  |||||||
    211  |||||||
    212  |||||||
    213  |||||||
    214  |||||||
    215  |||||||
    216  |||||||
    217  |||||||
    218  |||||||
    219  |||||||
    220  |||||||
    221  |||||||
    222  |||||||
    223  |||||||
    224  |||||||
    225  |||||||
    226  |||||||
    227  |||||||
    228  |||||||
    229  |||||||
    230  |||||||
    231  |||||||
    232  |||||||
    233  |||||||
    234  |||||||
    235  |||||||
    236  |||||||
    237  |||||||
    238  |||||||
    239  |||||||
    240  |||||||
    241  |||||||
    242  |||||||
    243  |||||||
    244  |||||||
    245  |||||||
    246  |||||||
    247  |||||||
    248  |||||||
    249  |||||||
    250  |||||||
    251  |||||||
    252  |||||||
    253  |||||||
    254  |||||||
    255  |||||||
    256  |||||||
    257  |||||||
    258  |||||||
    259  |||||||
    260  |||||||
    261  |||||||
    262  |||||||
    263  |||||||
    264  |||||||
    265  |||||||
    266  |||||||
    267  |||||||
    268  |||||||
    269  |||||||
    270  |||||||
    271  |||||||
    272  |||||||
    273  |||||||
    274  |||||||
    275  |||||||
    276  |||||||
    277  |||||||
    278  |||||||
    279  |||||||
    280  |||||||
    281  |||||||
    282  |||||||
    283  |||||||
    284  |||||||
    285  |||||||
    286  |||||||
    287  |||||||
    288  |||||||
    289  |||||||
    290  |||||||
    291  |||||||
    292  |||||||
    293  |||||||
    294  |||||||
    295  |||||||
    296  |||||||
    297  |||||||
    298  |||||||
    299  |||||||
    300  |||||||
    301  |||||||
    302  |||||||
    303  |||||||
    304  |||||||
    305  |||||||
    306  |||||||
    307  |||||||
    308  |||||||
    309  |||||||
    310  |||||||
    311  |||||||
    312  |||||||
    313  |||||||
    314  |||||||
    315  |||||||
    316  |||||||
    317  |||||||
    318  |||||||
    319  |||||||
    320  |||||||
    321  |||||||
    322  |||||||
    323  |||||||
    324  |||||||
    325  |||||||
    326  |||||||
    327  |||||||
    328  |||||||
    329  |||||||
    330  |||||||
    331  |||||||
    332  |||||||
    333  |||||||
    334  |||||||
    335  |||||||
    336  |||||||
    337  |||||||
    338  |||||||
    339  |||||||
    340  |||||||
    341  |||||||
    342  |||||||
    343  |||||||
    344  |||||||
    345  |||||||
    346  |||||||
    347  |||||||
    348  |||||||
    349  |||||||
    350  |||||||
    351  |||||||
    352  |||||||
    353  |||||||
    354  |||||||
    355  |||||||
    356  |||||||
    357  |||||||
    358  |||||||
    359  |||||||
    360  |||||||
    361  |||||||
    362  |||||||
    363  |||||||
    364  |||||||
    365  |||||||
    366  |||||||
    367  |||||||
    368  |||||||
    369  |||||||
    370  |||||||
    371  |||||||
    372  |||||||
    373  |||||||
    374  |||||||
    375  |||||||
    376  |||||||
    377  |||||||
    378  |||||||
    379  |||||||
    380  |||||||
    381  |||||||
    382  |||||||
    383  |||||||
    384  |||||||
    385  |||||||
    386  |||||||
    387  |||||||
    388  |||||||
    389  |||||||
    390  |||||||
    391  |||||||
    392  |||||||
    393  |||||||
    394  |||||||
    395  |||||||
    396  |||||||
    397  |||||||
    398  |||||||
    399  |||||||
    400  |||||||
    401  |||||||
    402  |||||||
    403
```

Search completed: April 28, 2003, 21:44:09  
Job time : 357 secs